

University of Richmond

UR Scholarship Repository

Honors Theses

Student Research

4-12-2021

Speech and Sovereignty: A Kantian Defense of Freedom of Expression

Alec Greven
University of Richmond

Follow this and additional works at: <https://scholarship.richmond.edu/honors-theses>



Part of the [Leadership Studies Commons](#)

Recommended Citation

Greven, Alec, "Speech and Sovereignty: A Kantian Defense of Freedom of Expression" (2021). *Honors Theses*. 1579.

<https://scholarship.richmond.edu/honors-theses/1579>

This Thesis is brought to you for free and open access by the Student Research at UR Scholarship Repository. It has been accepted for inclusion in Honors Theses by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.



SPEECH AND SOVEREIGNTY

A Kantian Defense of Freedom of Expression

By Alec Greven

Honors Thesis
Submitted to:
The Jepson School of Leadership
University of Richmond, VA
April 12, 2021

Project Advisor: Dr. Jessica Flanigan
Committee Members
Dr. Terry Price
Dr. David Lefkowitz

Abstract**Speech and Sovereignty: A Kantian Defense of Freedom of Expression**

By Alec Greven

Committee Members: Dr. Jessica Flanigan, Dr. Terry Price, Dr. David Lefkowitz

This thesis critically examines the moral foundations of free expression and offers a framework for evaluating morally justifiable forms of censorship. This investigation has three parts. The first section argues that rational considerations constrain how moral principles for censorship can be structured methodologically. It concludes that moral principles must be universally coherent and consistently applied. The second section considers several existing justifications for censorship that fall short of these methodological requirements and arbitrarily apply extensionally inadequate moral principles. To be rational, these approaches must either abandon these inconsistent justifications or commit to more consistently authoritarian moral principles. The third section outlines several methodologically consistent principles and ultimately defends the liberal model of free expression as the most plausible censorship principle that institutional leaders should adopt. This model restricts the censorship of speech in all cases except where censorship is necessary to defend the autonomy of others from rights violations.

Signature Page for Leadership Studies Honors Thesis

**Speech and Sovereignty:
A Kantian Defense of Freedom of Expression**

Thesis presented

by

Alec Greven

This is to certify that the thesis prepared by *Alec Greven* has been approved by his/her committee as satisfactory completion of the thesis requirement to earn honors in leadership studies.

Approved as to style and content by:



Dr. Jessica Flanigan, Chair

Dr. Terry L. Price, Member

Dr. David Lefkowitz, Member

Dr. Crystal L. Hoyt, Associate Dean for Academic Affairs
Jepson School of Leadership Studies

“Until every soul is freely permitted to investigate every book, and creed, and dogma for itself, the world cannot be free. Mankind will be enslaved until there is mental grandeur enough to allow each man to his thought and say.” –Robert Ingersoll

Preface

There are few things as distinctly important to humans as the ability to speak to others in a language that they will understand. Speech allows us to cooperate, to express our deepest feelings, and gives structure to how we perceive the world. This enormous value that speech has for us implies that there is also something of deep moral value in defending the ability of people to speak freely. This thesis investigates the moral foundations of free expression and proposes a framework for evaluating morally justifiable forms of censorship. I argue that the only morally justifiable form of censorship is the liberal model of censorship which restricts the censorship of speech in all cases except where censorship is necessary to defend the autonomy of others from unjustified interference.

The most famous defense of the liberal model of free expression against most forms of government censorship was the book *On Liberty* by John Stuart Mill.¹ He argued that the government should refrain from censorship except in cases of harm to others in order to defend individual freedom and preserve the pursuit of truth. Mill’s argument against censorship was ultimately based in his belief that a liberal society would lead to the most social utility on balance and that individuals have the moral obligation to maximize utility. Therefore, free expression for Mill is contingently defended so long as it is demonstrated that it actually promotes more utility. The deluge of offensive speech, hate speech, misinformation, and blatantly false speech

¹ John Stuart Mill, *On Liberty*, ed. David Bromwich and George Kateb (Yale University Press, 2003).

that has entered public discourse has led many people to be skeptical of the liberal project and whether or not it produces more social benefits on balance. Additionally, Mill's claim that speech can be regulated in cases where it causes harm to others needs more elaboration. After all, telling people certain true things can harm them in some cases and certain kinds of ultimately helpful criticism can cause pain. There needs to be a clearer articulation of the specific circumstances where the government is morally permitted to intervene and censor speech on the basis of harm. Thus, Mill's defense of free expression, while brilliant, has two major shortcomings because it defends speech on a contingent basis that may no longer hold in modern times and because it fails to offer a clear framework for which kinds of harms are enforceable and warrant censorship.

The aim of this thesis is to address these shortcomings and rest Mill's liberal defense of free expression on sturdier foundations. I offer a Kantian defense of freedom of expression which grounds the moral value of free expression in the sovereign autonomy of individuals. This liberal model of free expression says that censorship is only morally justified to defend the equal freedom of rational agents to exercise their autonomy. I defend autonomy as the guiding moral value to resolve questions of censorship. Protecting the autonomy of others is not a contingent defense of free expression like Mill's utilitarian defense is because free expression is itself an exercise of rational agency so it is constitutive and inherent to the value of autonomy. Unjustified violations of autonomy are the only kinds of harm that should fall under Mill's Harm Principle and should allow the government to engage in censorship. In this thesis I will elaborate on what makes certain types of speech unjustified violations of autonomy and therefore what are the classes of cases where censorship can be exercised. My proposed account of the liberal model of free expression resolves the

issues that Mill's argument faces because it makes it so free speech is no longer defended contingently and provides a more robust specification of Mill's Harm Principle. Mill would not accept my modifications to his theory because he is a utilitarian and my account is Kantian. Nonetheless, I offer this conclusion as a more plausible moral theory to make the case against most forms of government censorship.

This investigation has three parts. The first section discusses the source of why rational agents have moral obligations. I advance the claim that moral obligations flow from the need for rational agents to deliberate and make choices in a manner that is intelligible so that their agency can be maintained over time. Morality is ultimately a system of rationality where the structure of reason imposes constraints on what rational agents are able to do. I argue that the structure of reason is grounded in two core principles. The first is the Principle of Universality which states that the objective state of intelligible reality must be true and universal for all rational agents. Thus, moral laws cannot apply just to a particular individual or culture and must instead be objective in character. The second principle is the Principle of Coherence which states that all valid principles must cohere and cannot contradict one another. These principles ultimately set constraints on what can be a valid principle for moral agents methodologically. Any morally justifiable censorship principle must meet these standards.

There are several proposed theories of censorship which fail to satisfy these principles of structural rationality. A hypocritical censor is someone that holds individuals to unequal standards and fails to act in a universally coherent manner. Rational agents have the duty to avoid hypocrisy because they have the duty to act rationally and hypocrisy is not rational. I offer censorship and lying as cases that are

generally wrong because they are normally hypocritical. Rational agents are normally committed to preserving an effective system of communication but both lying and censorship tend to subvert this system of communication. If someone censors according to a principle that is not universally coherent while practically committing themselves to a system of effective communication then they are engaging in a contradiction. This contradiction is hypocritical because the person is holding themselves to a different standard than they hold others.

The second section of this thesis considers several existing justifications for censorship that fall short of the requirements demanded of rational agents. These are censors which advance themselves as “liberal” because they do not want to promote widespread and extensive government censorship of speech but they also want to regulate speech along the lines of cultural standards, offense, hate speech, and the promotion of democratic discourse. However, I argue that if these principles are extended in a nonarbitrary way then they will be incompatible with the practical commitment of preserving a liberal society that protects broad expressive freedoms. Therefore, to meet one’s rational duties, these hypocritical censors must resolve these inconsistent commitments by either abandoning these censorship principles or commit to more consistently authoritarian censorship principles.

The third and final section of this thesis discusses several consistent censors which all meet the methodological rational requirements outlined in the first chapter. Nonetheless, these censors are not equal in terms of their plausibility. Adopting various consistent censors comes with several unpalatable implications. I argue that the censorship model that we should adopt is one based on a principle which meets the methodological rational requirements and that is the most plausible. I point to the

liberal, utilitarian, radical egalitarian, maximal autonomy, and truth censors all as viable consistent censors. However, I point out that each of these consistent censors are committed to foundational commitments that are in irreconcilable tension with each other. One cannot rationally choose to adopt more than one of these consistent censors. I defend the liberal censor model as the model that is both a consistent censor and the one that is most plausible. I conclude this thesis by arguing that this framework of the moral foundations of free expression commits institutional leaders to certain rational requirements. Institutional leaders have the responsibility to censor speech according to a valid moral principle and a moral principle that is plausible. The liberal censorship model is the only model of the ones discussed that meets both of these requirements. Therefore, leaders should only engage in censorship of speech when it is necessary to defend the autonomy of others from rights violations in order to restore a state of equal freedom between rational agents.

Chapter 1: Morality as a System of Rationality

Introduction

Most people consider it wrong to interfere with the speech of others in at least many cases. This chapter will investigate moral obligations as it relates to free expression by considering the source of our moral obligations more generally. I will begin by arguing that humans have a duty to act in ways that are consistent with their rational nature. Acting consistently with one's rational nature requires that one acts on the basis of universally coherent principles. Therefore, our moral obligations must be formally structured in a way that makes our moral duties universalizable and coherent with our other moral principles. This formal structure of rationality explains why censorship is generally wrong. Most acts of censorship cannot be formulated in a way that can be universalized and coherent with other principles of action.

The chapter will then discuss the similarities between lying and censorship. Both of these actions are viewed as generally impermissible. However, many think that there can be principled justifications to explain why these acts can be morally permissible in certain circumstances. I will argue that the same reasons that justify lying in certain circumstances will also justify when censorship can be permissible. The chapter will conclude by defending a particular substantive thesis about lying and censorship which states that these types of typically impermissible behaviors are only permitted in defense of oneself or others from unjustified interference. I refer to this substantive thesis as the liberal framework which asserts that it is permissible to interfere with others when they unjustifiably infringe on the autonomy of others. Autonomy is the core value that is defended on a principled basis universally for each

rational agent. As autonomy is the controlling value, appealing to it avoids a rational contradiction. Later chapters will expand on this substantive thesis in more detail.

Section I: The Source of Normative Obligations

Humans are agents. In addition to having experiences we are also able to reflect on our experiences and judge them. This capacity for reflection allows us the ability to deliberate on different ends and pursue them based on our values. In other words, reflection helps us determine what we ought to do. The exercise of agency is fundamentally about reflecting about experiences and incorporating these reflections into an overall framework of an agent's ends or values. The following argument outlines what I believe to be the source of rational obligations for humans:

1. Humans are agents
2. Agents ought to live intelligible lives
3. Rationality is the mechanism by which one's life and reality becomes intelligible
4. Therefore, humans ought to be rational

The first premise appears to me to be pretty straightforward. Agency is characterized by the ability to form higher order judgments. Both knowledge and abstract thinking requires a hierarchy of thought where someone makes judgments about judgements. In his essay "Freedom of the Will and the Concept of a Person," Harry Frankfurt explains that personhood is established through the structure of a person's will which has the "capacity for reflective self-evaluation that is manifested in the formation of second-order desires."² A second order desire is a desire about one's

² Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy* 68, no. 1 (1971): 7.

desires and is a demonstration of agency because it is a hierarchy of judgement about one's judgments. Gary Watson offers a diverging view claiming that persons are able to hold values which "consist in those principles and ends which he – in a cool and non-self-deceptive moment – articulates as definitive of the good, fulfilling, and defensible life"³ Once again, determining values is an exercise of agency because it involves a hierarchy of thought establishing a higher set of principles among a variety of possible ones. For the purposes of this essay we do not need to choose between these accounts because agency is a common thread. In both views of personhood, there is the ability for a human being to engage in higher order reflection about states of affairs in order to determine what has claims on a person in terms of what they should do. As the claim that humans are agents is not very controversial, I will proceed to the second premise.

The second premise is that agents ought to live intelligible lives. Richard Feldman argues that there "are oughts that result from one's playing a certain role or having a certain position" and that oughts are predicated on good performance.⁴ For instance, if someone is a teacher then they ought to teach because being a teacher is predicated on the condition of a teacher teaching. Similarly, if a human being is an agent then they ought to act agentically, otherwise they are not engaging in good performance of what it means to be human. Returning to Frankfurt and Watson, a person is defined by either their second order desires or their values. Having either a second order desire or values requires a person to reflect and form judgements about their beliefs. If someone is not acting intelligently and reflecting and forming judgments about their beliefs then they are forsaking the defining element of their

³ Gary Watson, "Free Agency," *The Journal of Philosophy* 72, no. 8 (1975): 215.

⁴ Richard Feldman, "The Ethics of Belief," *Philosophy and Phenomenological Research* 60, no. 3 (2000): 676.

personhood. So, an intelligent being ought to live an intelligible life because an intelligent being is inescapably placed in a role or position which requires a certain kind of performance. If we are to play a role of what it means to be a person then we ought to play this role well. We would be incredulous of a teacher who decided they would maintain their position and teach poorly. Similarly, we would think that an intelligent being is not behaving as they ought to if they are not living an intelligible life.⁵

One may wonder at this point why agents are committed to living intelligible lives. I have claimed that the exercise of agency requires reflection and fitting one's reflection into an overall hierarchy of values and ends for an agent. For example, a person can see weights as not simply something to lift. Rather, their reflections over the weights allow them to see the weights as part of an overall end the agent holds of becoming healthy. A planning agent sets ends and setting ends requires an agent to form evaluations about how to realize their long term plans. Squirrels, very young children, and those with severe dementia lack this reflective capacity to form long term values and ends and that is why we hold them to different standards of what we expect of them. Unified agency recommends certain forms of action which maintain that agency across time. If an agent's agency is unified through a value such as health then a unified agent needs to pursue courses of action that maintain that unity. If an agent fails to live an intelligible life then they lack unified agency because unified agency must be bound together through intelligible patterns that reflect an agent's long term values and ends. An agent who commits herself to contradictory ends will

⁵ Some might wonder if all positional roles imply an obligation to perform that role well. Should thieves steal because that is their positional role? Usually no. Thieves do not have this obligation unless there is the obligation to become a thief in the first place. One cannot escape being a rational agent like a thief can escape being a thief so there is a positional obligation that comes with rational agency that does not apply in other cases.

strain her agency because such commitments cannot be simultaneously held. Insofar as we view ourselves as unified agents, we have the responsibility to make our lives intelligible in order to maintain our agency over time.

The third premise in my argument is that rationality is the mechanism by which one's life and reality becomes intelligible. Rationality and intelligibility are intimately connected because rationality unifies agency. I define a reason as a consideration in favor of discerning intelligible reality. I define Reason with a capital R as the actual state of intelligible reality. Every mechanism has to be aimed at furthering some end. The process of responding to reasons is a mechanism that is aimed as discerning Reason, or the state of intelligible reality. Intelligible reality is understood as objective features of the world in which humans live. The laws of physics, material objects, and our bodies are all features of what objectively exists. Our process of reasoning assists us in making determinations about the nature of the world as it is. When we make judgements after seeing objects fall our hand waving across our face we are ultimately reasoning about what is real.

In addition to the patterns of the physical world, there is also a state of intelligible reality with respect to morality. My claim is that there is an objective state of morality, just like there is an objective state of reality with respect to material objects and natural laws. I define objective morality as the state of intelligible reality with respect to right conduct. Therefore, when we are reasoning about our moral obligations we are searching for considerations in favor of discerning intelligible reality with respect to right conduct. If morality is in fact an objective state of affairs in our reality then it completely aligns with Reason which is the actual state of reality. When we reason about morality our search for morality and Reason is one in the same.

Moral obligations flow from the need for rational agents to deliberate and make choices in a manner that is intelligible. If life's choices are to be made intelligible then there must be governing principles that are able to impose form on what we should do in the world. Christine Korsgaard explains that "Normative concepts exist because human beings have normative problems. And we have normative problems because we are self-conscious rational animals, capable of reflection about what we ought to believe and do."⁶ Our normative problems arise because humans are intelligent and have the ability to assess our judgements relating to certain options being better than others. The nature of our agency means we are committed to forming judgements about our judgments and this requires us to stand by certain desires and values over others. Jose Ortega y Gasset once wrote that "To live is to find ourselves fatally obliged to exercise our liberty to decide what we are going to be in this world. Not for a single moment is our activity of decision allowed to rest. Even when in desperation we abandon ourselves to whatever may happen, we have decided to not decide."⁷ The process of deliberation runs into an issue in that we are constantly searching for why our reasons are justified by other reasons and what in turn justifies those reasons. Unless there is a stopping point, our reasoning will lead to an infinite regress that lacks a stopping point. Korsgaard notes that "The realist move is to bring this regress to an end by fiat: he declares that some things are *intrinsically* normative."⁸ Thomas Hill explains that deliberative agents are searching for "the ultimate ends that, for us, justify particular choices but themselves require no further justification."⁹

⁶ Christine Korsgaard, *The Sources of Normativity* (Cambridge University Press, 1996), 46.

⁷ Jose Ortega Gasset, *The Revolt of the Masses* (W.W. Norton & Company, 1994).

⁸ Korsgaard, *The Sources of Normativity*, 33.

⁹ Thomas Hill, "Pains and Projects: Justifying to Oneself," in *Autonomy and Self-Respect* (Cambridge University Press, 1991), 176.

The need to make one's choices justifiable necessarily commits a deliberative agent to the normativity of rationality. Korsgaard states that for a realist the "notion of normativity or authority is an irreducible one."¹⁰ The reason why the obligation is irreducible is because it is based in Reason itself. Ultimately, if deliberation yields a conclusion that is completely justifiable then it binds rational agents because an obligation of a rational agent is to follow rational dictates. Barbara Herman explains that "The *result* of deliberation is obligation. The practical necessity that is the core of Kant's view of morality arrives as the agent determines which of the grounds of obligation present binds to duty."¹¹ Remember, human beings are agentic beings so the nature of our lives forces us to make our lives intelligible. Reason reflects the state of intelligible reality so we are obligated to be rational out of necessity of how our being is constituted. Kant believes that through deliberation we bind ourselves to duty by determining what we have most reason to do. What we ought to do is derivative from what we are obligated to do and what we are obligated to do is derivative of what we have most reason to do. Therefore, we have a duty to be rational because rationality is the source of duty.

Reason subsumes all other potential sources of normativity because with all other alleged sources of normativity we search out reasons for why we think that something is actually a source of normativity. We justify eating healthy in the name of health, we justify health in the name of happiness, and we even find justification in making ourselves happy in order to make our lives intelligible to ourselves. All reasoning works ultimately to terminate in a state of intelligibility for a deliberative agent. Self-justification through reasoning is ultimately the only intrinsic source of

¹⁰ Korsgaard, *The Sources of Normativity*, 30.

¹¹ Barbara Herman, *The Practice of Moral Judgment* (Harvard University Press, 1993), 168.

normativity. It is only until we view something as reasonable will we, as rational agents, be able to accept it as a source of normativity for us. The only thing that can most closely approximate reasonability for us is reason itself so it makes sense that ultimately reason would be the only viable source of normativity. Thus, I believe that reason is the one distinct source of normativity.

Someone may ask why we should feel obligated at all to follow reasons' dictates. In other words, why should we be rational or care about acting rationally? The answer lies in the essence of who we are. We unavoidably have to live as rational agents and consequently rationality provides a governing structure for how we engage, interpret, and interact with the world. If we choose to live then we are making a conscious choice to engage, interpret, and interact with the world. This choice forces us to engage with the governing structure of rationality and give it credence. Rationality never loses its normative grip on us. As the end of rationality is Reason itself there can be no conflicting judgements from differing sources of normativity. In his book *The Importance of Being Rational*, Errol Lord thinks that "rationality consists in correctly responding to the objective normative reasons that an agent possesses."¹² There can only be conflicts for a rational agent in terms of differences about what are correct objective normative reasons in a relevant situation. Rationality as the source of normativity tells us to always pursue the most rational course of action and that is what grounds our normative obligations. Kant emphasized that "the content of the law comes from reason, not from anything special about you, or *your* reason, or even *human* reason, but from reason as such."¹³ This claim is important because it grounds

¹² Hallyard Lillehammer, "The Importance of Being Rational," *Notre Dame Philosophical Reviews*, April 2019, <https://ndpr.nd.edu/news/the-importance-of-being-rational/>.

¹³ Korsgaard, *The Sources of Normativity*, 174.

morality in an independent objective standard that abstracts away from a particular culture or individual. As morality is not particular to any individual or culture, we need an independent objective force that can provide the foundation of our moral obligations. The structure of Reason readily provides this foundation.

An objection might be made at this point that says that there is not an objective character to morality. This argument says that morality can be particular instead of objective and relate completely to either an individual or culture. This approach is flawed because it is contrary to basic principles of reasoning. Every culture and individual is committed to rational coherence when it comes to physics and mathematics. We do not accept John's claim that $2+2=5$ are "his mathematics" and we have ours. All rational agents are committed to finding coherent principles of action that make their lives intelligible. Agents are committed to finding coherent principles when it comes to physics and mathematics. Coherence must also be sought when it comes to determining ethical principles.¹⁴

The purpose of this section is to argue that humans have a moral obligation to be rational. I argued that humans are by nature agentic creatures and we have the ability to form judgements about our judgements and reflect on our actions and desires. As agentic creatures, we have obligations to make our lives intelligible. Our capacity for reflection makes it impossible for a rational creature to simultaneously hold the beliefs that X exists and X does not exist. Believing simultaneously that X exists and X does not exist is unintelligible and our ability to reason is what forces us to see that the two claims existing is unintelligible. Our rationality is the mechanism that makes the world intelligible because our reasoning is aimed at discerning the

¹⁴ The claim that ethics is relative unlike physics or mathematics will be addressed in the second chapter.

state of intelligible reality. Reason with a capital R reflects the actual state of intelligible reality and is the end state of reasoning. When agents deliberate they are fundamentally searching for reasons that justify themselves with reference to the actual state of intelligible reality. We may fail in our deliberations but the nature of our lives makes it so we cannot avoid deliberation. Thus, if we ought to make our lives intelligible and reasoning is the only way to do this then we ought to be rational. What we ought to do sets a framework for what we are obligated to do. Morality reflects our obligations. If we have the obligation to be rational then we have the moral obligation to be rational. The main conclusions of this section are that Reason is the source of normativity for rational agents and that intelligent creatures have a moral obligation to be rational.

Section II: The Character of Reason

While I have argued that Reason is the source of normative moral obligation, I have not yet touched on what the form of Reason looks like yet. This section will lay out the character of Reason. I will argue that the character of Reason takes the form of universally coherent principles. Any source of moral obligation must conform to the character of Reason and this will give us insights into how we are able to frame our moral obligations and justifications. I will advance two principles that reflect the form of Reason which are derived from Kantian arguments. These principles are:

- The Principle of Universality
- The Principle of Coherence

Before I dive into the principles, I want to first outline the standards of what constitute the requirements of Reason. Recall that Reason is defined as the state of intelligible reality. If a principle is to approximate intelligible reality then it must be accessible to all intelligent agents. Intelligent agents make reality intelligible through the mechanism of reason so the same use of reason for all intelligent agents needs to yield the same conclusion. Moreover, intelligible reality cannot be contradictory because anything contradictory is unintelligible. Both of the principles are aimed at capturing these formal constraints of Reason and these constraints set limits on how rational agents are able to conceive of their moral obligations.

I will call the first principle of Reason the principle of universality. This principle states that any conclusion accurately reflecting the state of intelligibility must be universally and conclusively reached by all equally rational agents who impartially weigh the considerations at issue. The objective state of intelligible reality must be true for all rational agents. Kant points out that “as morality serves as a law for us only because we are rational beings, it must also hold for all rational beings.”¹⁵ For instance, laws of mathematics and physics are universal laws that have universal force in the world. Take Planck’s Constant or the Pythagorean Theorem. These laws reflect objective relationships in the world and are a representation of universal laws. Similarly, objective moral laws need to be binding on all rational agents. Kant explains that the principles of reason are a priori principles so they “cannot be obtained by abstraction from any empirical, and therefore merely contingent, knowledge.”¹⁶ He notes that a command of reason is an imperative and that objective principles of

¹⁵ Immanuel Kant, *Groundwork of the Metaphysic of Morals*, trans. Thomas Abbott (Digireads.com Publishing, 2017), 51.

¹⁶ Kant, 23.

reason are based in “principles which are valid for every rational being as such.”¹⁷ The law of gravity, Planck’s Constant, and the Pythagorean Theorem are all valid principles for every rational being and are constitutive principles of Reason. For a moral law to be a valid principle of Reason then it must be a valid principle for every rational being, or universal in nature.

Kant formulates the requirement of the principle of universality through the Categorical Imperative. An imperative is an objective principle that is binding for rational agents. If morality is to have imperatives then they must be categorical and bind all rational agents equally. Kant says the Categorical Imperative has the formulation of “Act only on that maxim whereby thou canst at the same time will that it should become a universal law.”¹⁸ The Categorical Imperative is an expression of principle of universality because if a principle is to satisfy the formulation of the Categorical Imperative then it must be universally and conclusively willed by all rational agents who impartially weigh the considerations at issue. Terry Price notes that because “morality does not rest on any particular reason for action, no change in circumstances could undermine the authority of the categorical imperative.”¹⁹ The Categorical Imperative expresses a form for principles but itself has no content or relies on no facts on the ground. This is why Kant states that “Empirical principles are wholly incapable of serving as a foundation for moral laws.”²⁰

At this point, I am not taking a stand on any particular moral philosophy. For example, diverse philosophies like ethical egoism and utilitarianism can both satisfy

¹⁷ Kant, 25.

¹⁸ Kant, 31.

¹⁹ Terry Price, *Leadership Ethics: An Introduction* (Cambridge University Press, 2008), 52.

²⁰ Kant, *Groundwork of the Metaphysic of Morals*, 47.

the principle of the Categorical Imperative. The willing of the principle that all rational agents should promote their self-interest or all rational agents should maximize utility both in theory satisfy the formula of the Categorical Imperative. They express universal principles with respect to every rational agent. What this principle does is establishes a groundwork framework that any valid moral principle must meet.

The second principle that reflects the form of Reason is the principle of coherence. This principle asserts that all valid principles must cohere and cannot contradict one another. Any state of intelligible reality cannot be contradictory because contradiction renders a state of affairs unintelligible. It is unintelligible for an agent to simultaneously believe X and not-X. Moreover, reality cannot simultaneously consist of X and not-X. Therefore, a valid principle of Reason is to have principles that cohere, there must either be X or not-X. Moral principles must be coherent in order to be binding on rational agents. It is unintelligible to have two moral principles that one “cannot kill humans” and “I can kill that human named Bob.” Both of these principles considered simultaneously contradict each other and are incoherent taken together.²¹

The principle of coherence places a second major constraint on the form of valid moral principles. Established principles cannot contradict other established principles. Kant explains that “all maxims ought by their own legislation to harmonize with a possible kingdom of ends as with a kingdom of nature.”²² When a rational being sets reasonable ends then they must not conflict with the reasonable ends of other

²¹ A lot of behavior will be intelligible according to this standard. As we shall see in the third chapter, there are various diverse approaches to censorship that all can be coherent. But the principle of coherence still places important constraints on the framing of moral principles.

²² Kant, *Groundwork of the Metaphysic of Morals*, 43.

rational agents or not conflict with any laws of nature. Kant's kingdom of ends is ultimately a vision of perfect coherence between all ends of intelligible reality.

Section III: Authenticity and Hypocrisy

When the principles of universality and coherence are combined they provide a general framework for evaluating the validity of moral principles. Any posited moral principle that is not universal for all rational agents or contradicts with another moral principle fails to establish itself according to a rational form and should be rejected. These principles also limit agents morally from acting on the basis of certain principles that are incoherent or cannot be universalized. An implication of this conclusion is a duty of authenticity for rational agents. I define an authentic agent as an agent whose expressive behavior aligns with the agent's beliefs. The duty to be authentic flows from an agent's duty to be coherent. If an agent believes in a particular moral principle then they act irrationally when they act in a way that is not in accord with their beliefs, either their actions or beliefs are wrong. For instance, if an agent believes that stealing is an immoral act and then steals she is acting inauthentically because her beliefs do not align with her actions. Authenticity can be reestablished by either changing her beliefs about the permissibility of theft or changing her actions. I argue that all rational duties imply a duty to be authentic which flows from a rational agent's duty to be coherent in both their actions and beliefs.

The duty to be authentic explains the wrongness of hypocrisy. Jay Wallace explains that hypocrisy typically entails "an inconstancy in a person's attitudes and behavior. People whose outward behavior does not comply with their inner convictions seem to

be at odds with themselves.”²³ Hypocrisy is the defining act of inauthenticity because it reflects an incoherence between an agent’s expressed actions and beliefs. Hypocrisy involves attitudes of moral blame for others when the same agent is blameworthy in the same way. Wallace emphasizes we act objectionably as hypocrites when we blame others “without subjecting your own attitudes and behavior to critical assessment, and bringing them into harmony with your current reactions to the attitudes and behavior of others.”²⁴ Wallace thinks that we act wrongly as hypocrites because we fail to respect the equal moral standing of others by attaching “to my interests greater importance than it ascribes to yours.”²⁵ The duties of rationality require that we act according to maxims that we can will as universal laws. Hypocrites do not act in a manner that can be universalized because they apply unequal standards of blame and fail to treat others as moral equals. Rational agents have the duty to avoid hypocrisy and act according to universally coherent principles. Therefore, any valid moral theory cannot allow for hypocritical blame or for agents to act inauthentically.

I will now explain why lying and censorship are generally impermissible. Performing these actions usually generates a rational contradiction. If a liar or censor commits a rational contradiction, they act hypocritically by applying unequal standards between themselves and others. Lying and censorship are usually wrong because these actions fail to respect the equal moral standing of others due to the application of unequal standards that prioritizes the interests of some over others on the basis of rationally incoherent standards.

²³ Jay Wallace, “Hypocrisy, Moral Address, and the Equal Standing of Persons,” *Philosophy and Public Affairs* 38, no. 4 (2010): 309.

²⁴ Wallace, 326.

²⁵ Wallace, 328.

Section IV: The Wrong of Lying and Censorship

We are now able to tie the threads of this moral theory together and use these framework to explain the general wrong of censorship. I will argue that censorship is usually wrong because the censor commits herself to a kind of practical contradiction which violates the standards of rationality central to any kind of morally permissible action. Thus, the censor generally acts impermissibly. In order to explain how this practical contradiction works I will appeal to the generally immoral act of lying. The wrongness of lying can be explained by lying generate a practical contradiction that commits an agent to irrational action. However, there are principles that can be appealed to in order to justify certain exceptions to the general rule that one ought not lie. I will outline these kinds of exceptions and point to similar exceptions that can be appealed to in order to justify certain kinds of censorship.

Christine Korsgaard says that certain actions are contrary to our moral duties if they commit an agent to a practical contradiction. She explains that “If universalization would destroy the connection between action and purpose, the purpose is not a sufficient reason for the action.”²⁶ Take the case of someone cutting in line. Someone cuts in line in order to gain an advantage of waiting less. However, if everyone cutting the line were to be universalized then the advantage someone would gain from the action would be frustrated; there would be no line to cut. The person who cuts the line makes them an exception to a general rule that people should organize themselves into a line. Line cutting functions as a kind of practical contradiction that Korsgaard explains where universalizing line cutting would break

²⁶ Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge University Press, 1996), 102.

the connection between the action and its intended purpose. The agent who cuts the line wills two incoherent things: that there is a line and that they can cut the line.

Rational agents need to be able to will commitments that do not undermine their other commitments. Rational agents are thereby limited in the actions that they are able to consider morally permissible actions because certain actions will result in a practical contradiction of will.

The reason why no rational agent can permissibly commit themselves to a practical contradiction is because it is irrational to do so. When an agent engages in a practical contradiction they see themselves as an exception to a rule they accept. An agent who commits a practical contradiction acts unintelligibly because they commit themselves to contradictory principles. This action threatens the unity of an agent over time because agency cannot be unified according to contradictory principles. An agent is defined by their reflective evaluations in light of values and ends the agent sets for themselves. However, if an agent has contradictory values or ends then this fragments agency. It is impossible for an agent to simultaneously commit themselves to the principles X and not-X. Trying to will both principles simultaneously is a practical contradiction and is wrong because it fragments the unity of a person's agency.

Those that commit practical contradictions are hypocrites through the application of unequal standards. Hypocritical action is wrong because it denies the equal moral standing of persons. Hypocrisy is itself a practical contradiction by applying a standard that should be applied to all while simultaneously denying the application of the principle to oneself. Individuals have moral standing in light of their rational agency. However, if a person acts hypocritically they engage in a practical contradiction by violating the standards of rationality that demand acting on the basis

of universally coherent principles. Whenever an agent fails to interact with others on the basis of universally coherent principles then they undercut their own moral standing which is derived from their rational agency. An agent cannot simultaneously will that they have moral standing while denying the moral standing of others. Therefore, hypocritical action is wrong because it involves incoherence on the part of the agent and fails to treat others as though they have equal moral standing.

I will now outline the value of communication. The capacity to effectively communicate with others is crucial for an agent to realize their distinct ends, projects, and values. All agents need to will a world in which the value of communication is preserved in order to realize their ends. Lying and censorship are two actions that subvert the value of communication. Thus, engaging in lying and censorship is usually a hypocritical action that commits an agent to a practical contradiction. It simultaneously commits an agent to a principle that the value of communication in the world should be preserved while performing actions that subvert the value of communication. If everyone lied and censored at will then the structure of communication that the agent is practically committed to would collapse. Therefore, the liar or censor makes themselves an exception to a rule which is hypocritical and fails to respect the unity of their agency and treat others with equal moral standing.

The Value of Communication

The effectiveness of human communication is one of the central reasons humans have advanced so much as a species. Communication enables collaboration and enables people to set plans and realize long term projects. Seana Shiffrin emphasizes that we need mutual access to the minds of others in order to live in

communities and realize our shared moral projects.²⁷ Helping others, coordinating donations to charity, collaborating in mutual defense are all long term moral projects that are successful in virtue of effective communication. A breakdown in our ability to communicate threatens the preservation of some of our most fundamental moral values. Once again, this is not a consequentialist defense of communication, although these arguments certainly are strong consequentialist reasons to defend the value of communication. I am claiming that every agent is practically committed to the value of communication to realize their distinct ends and projects.²⁸ This line of argument explains why every rational agent should practically commit themselves to maintaining an effective system of communication.

A shared moral community relies heavily on the notions of praise and blame. We praise or blame others to the extent that they succeed or fail in meeting their moral responsibilities. Blame is a type of sanction we apply in order to determine how we interact with others. As blame is a type of sanction we should be cautious about applying this sanction only to those who deserve it. One of reasons we do this is because we want to avoid the unjustified sanction of others that comes with blame. We do not want to be blamed when we fulfill our moral responsibilities or are incapable of doing so. Our determination of moral responsibility is heavily dictated by the mental states of the actors who engage in them. We praise the person who helps the poor when they do so out of concern for the welfare of others. We would blame a person or withhold praise if we found out a person only helped the poor to make their image look good and had no concern for the welfare of the poor. Many people think

²⁷ Seana Shiffrin, *Speech Matters: On Lying, Morality, and the Law* (Princeton University Press, 2014), 13.

²⁸ Of course, it is possible to conceive of a person who does not practically will a world with effective communication. I have yet to meet such a person.

that moral responsibility turns on the reasons for which people perform actions and not simply the actions that they perform. Communication is an important window that we have into the mental states of others to help evaluate their reasons for action even if we are never able to fully know why someone acted. Likewise, an effective system of communication also allows us to communicate our mental states to avoid sanctions from unjustified blame. A breakdown in dialogue is dangerous to our moral community because we would lose a key source to help us determine moral responsibility and appropriately praise and blame others.

A third reason why communication is so important is that it allows individuals to cultivate their agency as moral thinkers. Thinking is incredibly difficult when done solely within the confines of one's own mind. Effective thinking involves expressing one's ideas by drawing them out into the open in order to evaluate them. Shiffrin explains that in order to formulate a complex world a thinker must have the "ability to externalize bits of one's mind, identify them as particulars, and then evaluate them."²⁹ We are unable to become good moral agents without thinking in complex ways. Insofar as limits on communication place limits on our ability to think in complex ways our moral agency is limited. Without the ability to communicate we lose the ability to express ourselves and imprint our identity on the world around us.

A fourth reason why communication is valuable is because it is a necessary condition to fulfill an agent's duty of authenticity. I previously outlined how rational agents have a duty of authenticity that is based in coherence. I defined an authentic agent as someone whose expressive behavior aligns with their beliefs. If an agent is limited in the extent of their communication then this could infringe on their ability to

²⁹ Shiffrin, *Speech Matters: On Lying, Morality, and the Law*, 90.

say what they think and believe. Preserving the ability to communicate is crucial so that an agent can maintain coherence between their expressive activity and their beliefs and meet their rational obligations. Authenticity and honest communication are very similar but can come apart. I defined authenticity as the state of alignment between an agent's behavior and beliefs. When a person is dishonest they act inauthentically by intentionally create a gap between what the agent believes and how they represent their beliefs to others. A dishonest person behaves inauthentically but my definition of authenticity allows for more inauthentic behavior beyond dishonesty.

This is only a rough sketch of only some of the reasons why agents should practically will a world of effective communication. Communication is an integral part of our daily lives. Indeed, a human being is unlikely to have a single day in their lives go by without communicating in some form or fashion with others. My aim in this section is to establish why every person should be practically committed to preserving communication. This practical commitment constrains the conduct of rational individuals. As I previously noted, rational agents have the obligation to avoid practical contradictions. Our actions cannot contradict our practical aims. We cannot act in a manner that subverts the type of behavior that we want to will to be universal law. Therefore, actions that limit the extent of communication in the world are wrong *prima facie* because they compromise our practical aim to preserve the value of communication in the world. An implication of this is that any agent who engages in communication-limiting behavior has the burden of justifying their actions without resorting to a practical contradiction in will. Deception and censorship are generally wrong in that they usually unjustifiably limit our communication structure. I will now turn to several accounts of when deception can be justified. These contingent defenses

of deception will illuminate some contingent defenses of when censorship may be justified.

Principled Defenses of Deception

Deception threatens the key values that I outlined relating to communication and is a *prima facie* wrong. First, deception can compromise the ability for people to communicate in order to protect their moral projects. For instance, collective self-defense from those who wish us harm is an important moral project. However, to collectively defend ourselves we must rely on what others say. Deception is a wrong because it undermines our ability to rely on the communication of others as truthful which can limit our ability to engage in collective moral projects like self-defense. Weakening our ability to rely on the mental states of others as truthful also threatens our ability to hold others morally responsible. When we lose trust in the speech of others our ability to engage in dialogue with others and form external representations of the world is damaged. Finally, when we deceive others we violate a duty of authenticity we owe to ourselves. Deception necessarily involves a subterfuge between what someone expresses and what they actually believe about the world. Thus, the deceptive agent is an inauthentic agent. For these reasons we can conclude that deception is a *prima facie* wrong.³⁰

I also believe that deception is a *pro tanto* wrong. What I mean by this is that deception is always wrong to at least some extent. I previously mentioned that deception involves a type of dishonesty where there is a disconnect between an agent's expressive activities and their beliefs. Thus, the deceptive person is always inauthentic

³⁰ The moral wrong of deception also helps to anchor why certain forms of false speech make the person who says them liable to be censored.

to at least some extent. There is a loss when an agent acts inauthentically because it threatens the coherence of an agent's identity. While deception is a *pro tanto* wrong that does not mean that it cannot be outweighed by other considerations. There is always a loss when an agent is deceptive but all things considered in certain circumstances it may be permissible for an agent to deceive. I will now consider a situation where I think it is justified all things considered to deceive.

While deception is a *prima facie* wrong, most moral theories would make a principled exception for lying in certain circumstances. One case is the famous murderer at the door where a murderer approaches you and asks you if your friend is inside your house. If you refuse to answer your silence will reveal your friend's location. If you tell the truth then the murderer will know your friend's location and proceed to murder him. The paradox of communication is that subverting most of our communication structures is a *prima facie* wrong while many moral theories would assert it is permissible, or even obligatory, to lie to the murderer at the door. The rational framework above constrains the manner in which moral theories can admit justifications for lying. Principled defenses of deception must be compatible with a universal principle that does not contradict any other moral principles. The paradox of communication can be resolved with a moral theory that explains why lying is wrong in most cases but there are cases like the murderer at the door where lying can be justified without committing a practical contradiction.

Kant thought that lying was never justified. Thus, he denies there is even a paradox where cases of deception like the murderer is permissible. Kant went so far as to say that even when confronted with a murderer at your door you still have the duty

not to lie even to protect the intended victim.³¹ The reason why Kant thought you could not lie to the murderer at the door is because you would be engaging in a practical contradiction by lying. Korsgaard explains that Kant thought lying was a practical contradiction because the efficacy of a lie “depends upon the fact that most people do not engage in them, and which therefore can only be performed by someone who makes an exception of himself.”³² A successful lie to the murderer requires in Kant’s view that your lie is accepted as true. However, if everyone’s lies were accepted as true then the whole fabric of our communication system would break down. An agent cannot be simultaneously committed to the whole fabric of our communication system and lying in this case so there is a duty to not lie in order to avoid a practical contradiction. Thus, Kant thinks you have a duty to never lie, even to the murderer at the door.

Kant’s approach passes the formal requirements of reasoning because he advances a universally coherent moral principle. Nonetheless, substantively his account is very unsatisfying. When evaluating a moral principle we should first make sure that it is formally valid and then make sure it is a principle we are willing to accept in terms of substance. Saying that an individual can never permissibly lie is a very strenuous requirement. The most damaging implication of this hard position is that it fails to allow for lying in defense of oneself and others. We can imagine cases where individuals are in a position where the use of deception is the only way to avert serious moral harms. If you tell the truth to the murderer at the door then an innocent human being will die. Moreover, if you stay silent in the face of the question of

³¹ Christine Korsgaard, “The Right to Lie: Kant on Dealing with Evil,” *Philosophy and Public Affairs* 15, no. 4 (1986): 326.

³² Korsgaard, 329.

whether the intended victim is in your house you indicate to the murderer where your friend is. Deception is the only strategy, short of physically confronting the murderer, to prevent an egregious moral harm. Thus, I conclude that Kant's view should not be accepted substantively even if it passes formal requirements of reasoning.³³

Christine Korsgaard deviates from Kant's strong position and thinks an agent can lie to the murderer at the door without engaging in a practical contradiction. She contends that the lie to the murderer is not a practical contradiction because universalizing the lie would not destroy the efficacy of the lie.³⁴ Korsgaard notes that while you know the circumstances of the situation the murderer does not know that you know the circumstances of the situation. He is unaware that you know he is intending to murder the victim. Therefore, universalizing the lie does not destroy the efficacy of the lie because there is an informational asymmetry between you and the murderer which makes it successful even when the lie is universalized. Korsgaard also thinks a person can permissibly lie to the murderer at the door in order to prevent themselves from being used as an instrument for evil.³⁵ She argues that because the murderer is trying to deceive you deceptive defensive action is justified so that one does not become an instrument for evil ends. Thus, deception would not be generally permitted in Korsgaard's view but would be permitted in contingent circumstances where a lie can be universalized without creating a practical contradiction and in order to defend oneself from being used as an instrument for evil.

³³ Utilitarianism is another theory that is formally sound but not substantively. The utilitarian has a principle that any action that serves to maximize expected utility is morally required. I will consider the utilitarian argument in more detail in the third chapter.

³⁴ Korsgaard, "The Right to Lie: Kant on Dealing with Evil," 340.

³⁵ Korsgaard, 340.

Christine Korsgaard's account is better than Kant's because it admits a principled basis by which an individual is able to engage in deception in self-defense. Her argument passes through the standards of formal reasoning because it is universally coherent. Korsgaard articulates a principle that can be universalized and does not conflict with other moral principles. I contend that any moral theory that does not allow for the use of deception in any circumstances to defend oneself from being used as an instrument for evil is unsound substantively. I think this because individuals are entitled as rational agents to a sphere of autonomy that others are unable to breach and this sphere can be defended. The weakness of Korsgaard's view is that it does not allow for the use of deception in defense of others. I will now turn to why that is the case. I think allowing for defense of others is crucial because individuals have key interests in virtue of their autonomy as rational agents. We must recognize that others are also entitled to the same sphere of autonomy that we hold ourselves entitled to. Certain actions that violate the autonomy of others fail to respect their rights. A substantively adequate moral theory is one that allows for the defense of individual autonomy.

Tamar Schapiro takes issue with Korsgaard's argument because it inadequately addresses the permissibility of using deception in defense of others. Schapiro accepts that Korsgaard's analysis works for the murderer at the door but would not work in a situation is not using deception. Schapiro imagines a case where a Nazi approaches you to ask about where you are hiding fugitives and is completely honest about his intentions.³⁶ Korsgaard's view would say it is impermissible to lie to the Nazi because he is not trying to deceive you and is honest. Schapiro thinks deception is usually

³⁶ Tamar Schapiro, "Kantian Rigorism and Mitigating Circumstances," *Ethics* 117, no. 1 (2006): 52.

wrong because it violates the autonomy of others and “amounts to a refusal to reciprocate within a scheme of shared thought and action.”³⁷ Communication necessarily involves more than one person and successful communication requires certain actions be performed by each person involved. For this reason, Schapiro says that communication is a type of “colegislation game” with mutual obligations. She argues that we can permissibly deceive the murderer and Nazi at the door because their actions demonstrate a commitment to not respect the autonomy of others which erodes the foundational value of honest communication.³⁸

The duty of honest communication is derived from our duty to respect the autonomy of others. Our duties to respect the autonomy of others are based in an aspect of colegislation and individuals can forfeit their entitlements to have their autonomy respected by failing to respect the autonomy of others. Deception is usually wrong because it interferes with the autonomy of others and manipulates an individual’s understanding of the world so that they are not able to operate as an autonomous author of their own ends. However, when a person makes the decision to violate the autonomy of others they forsake entitlements they hold in virtue of their autonomy. Such a decision makes an individual liable to interference in direct proportion to the degree they threaten the autonomy of others.

Schapiro’s account is an improvement over Korsgaard’s because it elaborates on a principled argument for why lying is generally wrong but we can permissibly lie to the murderer at the door in defense of ourselves and others. This account is focused on protecting individual liberty from interference. I will classify Shapiro’s explanation

³⁷ Schapiro, 50.

³⁸ Schapiro, 54.

that resolves the paradox of communication as the liberal communicative theory. This theory asserts that individuals have the responsibility to respect the autonomy of others and this can involve engaging in communication strategies that protect oneself and others from unjustified interference to their autonomy. The liberal communication theory asserts that most individuals have rights against interference against interference in virtue of their autonomy as rational agents. Kant argues that rational agents have the sovereign right to independence where no one else can claim authority over them when they respect the equal freedom that is granted to rational agents.³⁹ This right to independence grounds the right of autonomy. The theory passes through the standards of formal reason because it advances a universally coherent standard that defends a universal principle that allows for deception in defense of others. This is also the correct view substantively because it respects autonomy as a preeminent value. Agency involves the setting of ends and values and bringing our reflective experiences in accord with these values and ends. Autonomy crucially is a reflection of agency because autonomous action establishes an individual's ends and values as distinctively hers. Individuals are sovereign agents and their autonomy must be respected otherwise the equal moral standing of individuals would be threatened.

There is also not a practical contradiction when deception is allowed in defense of the autonomy of oneself or others. The foundational value of communication is that it promotes autonomy and allows individuals to realize their own projects and ends. An agent will only practically will a system of effective communication insofar as it contributes to the preservation of a justified autonomous sphere of action for individuals. An agent commits no practical contradiction by lying in defense of the

³⁹ Arthur Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy* (Harvard University Press, 2010), 35.

autonomy of oneself or others because the agent is willing the universally coherent means in the world to respect their autonomy. Such means will largely demand honest communication but will admit of certain principled exceptions.

In this essay I am using a moralized conception of interference which is bounded by the entitlements others have. The liberal account does motivate the question of what constitutes justified interference with others and what specific entitlements individuals possess in terms of their autonomous sphere of action. In the second chapter I will review several arguments for what sphere of autonomous agency an agent is entitled to. Before we do that we can turn to principled defenses of censorship. What we shall see is that the same arguments regarding justifications of lying also apply to communication strategies that justify censorship.

Principled Defenses of Censorship

In this section I contend that censorship faces the same paradox of communication that deception does. First, the same arguments for the value of speech counting as reasons to avoid deception are the same reasons that count for avoiding censorship and in some cases allowing it. The act of censorship interferes against the ability of others to communicate. As I define it, censorship differs from persuasion in that it does not involve changing the expression of speech through voluntary actions. Censorship uses force in some form to override the will of others to speak. Frustrating the communication of others is a *prima facie* wrong like deception because it threatens our communication structures which protect crucial values. Remember, agents practically will an open structure of communication because they:

- Allow us to collectively work with others to realize our moral aims

- Provide us access into the minds of others to assign moral responsibility
- Enable individuals to cultivate their autonomy and individual identity
- Preserve the ability of agents to fulfill their duties to act authentically

Censorship threatens all of these practical aims by limiting communication. Therefore, I argue that it is a *prima facie* wrong to censor the speech of others.

The foregoing argument does not say it is always wrong to censor. Rather, censorship must be justified on a principled basis. For a principle to be rationally supported it must be universal and not contradict any other principles. Censorship has the same paradox of communication that deception encounters. Most forms of censorship are viewed as morally problematic. However, there are certain types of censorship that most moral theories would say are justified. For instance, consider Justice Holmes' famous case where a man intentionally yells fire falsely in a crowded theater starting a panic where multiple people are seriously injured.⁴⁰ Like the murderer at the door, most moral theories are pressed with the paradox of why most censorship seems unjustified but censoring the individual from intentionally causing a dangerous panic in a theater by falsely shouting fire is justified. I will argue that a plausible substantive moral theory would justify censorship of the person who intentionally shouts fire in a crowded theater on a principled basis.

The utilitarian would solve this paradox in the same way they solve the communication paradox relating to deception. Their approach is to only censor when doing so maximizes utility. Censorship in general is wrong because it can promote dogmatic thinking, can threaten the search for truth, and erode the ability for

⁴⁰ *Schenck v. United States*, 1919.

individuals to build their identity which contributes to their happiness.⁴¹ Nonetheless, the utilitarian would not say that all forms of censorship fail to maximize utility. The theater case is a case where it would maximize utility to censor a person from falsely shouting fire in the theater. The panic would cause serious injuries which would erode utility. Moreover, the utility in the speech act of the individual in speaking is very low. The person is intentionally shouting fire falsely in a theater with the express purpose of starting a panic. They are intentionally asserting known falsehoods in a moment where individuals are not able to carefully deliberate and respond appropriately. For all these reasons the utility would say censorship in the theater case is justified because it maximizes utility.⁴²

Some may adopt an approach similar to Kant's views on deception with censorship and deny there is even a paradox. They would assert that censorship, like lying, is never justified and the person falsely shouting fire in the theater cannot be censored. The argument claims that there is not a principled distinction in which a censor is able to avoid general censorship but censor the person in the theater case. Like Kant's position with the murderer at the door, this approach seems excessively rigorous. Yet, it is a universally coherent approach that adheres to the rational principles above. Substantively, however, I will reject this theory for the same reasons why I reject Kant's universal prohibition against deception because the approach does not allow for defense of oneself or others.

⁴¹ The utilitarian case against censorship has received its most famous and eloquent defense in John Stuart Mill's book *On Liberty*.

⁴² Ultimately, I do not accept this line of reasoning on a substantive basis. Here I am outlining how utilitarian reasoning could be used to address the censorship dilemma. I will return to utilitarian arguments for censorship in the third chapter.

Others can go the route Christine Korsgaard takes with respect to censorship. They may argue that censorship could be narrowly universalized where an individual can protect themselves from being used as an instrument for evil. However, Korsgaard's argument only applies in more limited cases where an individual is able to protect themselves from being used for evil ends. This account invokes an ideal of self-defense but does not extend the strategic communication to defending others.⁴³ Recall that Korsgaard's argument is vulnerable in that it would allow deception to the murderer at the door but would not allow deception to the forthright Nazi who is not deceiving you but is planning on murdering others. I take issue with Korsgaard's reasoning being applied to the censorship case because it fails to adequately allow for censoring speech in defense of others. Adopting Korsgaard's approach would permit censorship if *you are in the theater* and censoring the intentional liar would defend yourself and prevent you being used as part of an evil end. However, imagine this case:

Theater Defense of Others Case

You become aware that an evil actor is about to intentionally and falsely yell fire in a crowded theater in order to start a panic and cause serious injury. You are not in the theater and would not be injured if the evil actor yelled fire. Thus, you personally are not being used as an instrument for evil. However, you have a big red button in front of you. If you press it the evil actor will be silenced against his will and be unable to carry out his plan.

⁴³ In fact, Korsgaard's account may not even allow for censorship in defense of oneself in the theater case. In this case there is no information asymmetry between you and the person shouting fire. Korsgaard's position may have to commit itself to the same position as the rigorous Kantian approach would in the theater censorship case.

Korsgaard's view of justified lying is troubling because it does not allow for lying in defense of others. Similarly, if we apply her principle justifying lying to the murderer at the door to the theater censorship cases it would permit censorship by someone in the theater but would not allow censorship on behalf of others in a different theater. The theater case in defense of others strikes me as a case where censorship is clearly justified. Any moral theory that does not justify censorship in the theater case is not a compelling moral theory. Therefore, substantively I do not accept Korsgaard's argument as it relates to the theater defense of others case.

Tamar Schapiro's discussion of justifications for deception provides an important way to solve the paradox of communication. Remember, we are looking for a moral theory that explains why censorship is wrong in most cases but justifies censorship in certain cases on the basis of defense of oneself and others. This is what I will refer to as the liberal defense of free expression because it rests on individual autonomy. Schapiro notes that communication is a game that requires at least two parties in order to be successful. The value of communication is derived from respect parties have for the autonomy of others as rational agents.⁴⁴ When a person fails to respect the autonomy of others then this assaults the basis of why we view censorship as a fundamental wrong. When the evil actor plans to intentionally yell fire in a crowded theater to cause a panic she is compromising the autonomy of others. We have a duty to usually not censor in virtue of an agent's status as an autonomous rational agent. But when the evil actor seeks to violate the autonomy of other she is forfeiting her entitlement against certain kinds of censorship. She voluntarily decides

⁴⁴ Schapiro, "Kantian Rigorism and Mitigating Circumstances," 54.

to breach the standards of the co-legislation communication game and this forfeits her entitlement to be respected in certain ways.

We can see the justifications for lying to the murderer at the door neatly align with the same justifications for censoring speech in the theater case. Both examples involve a person who is infringing on the autonomy of others and the actions present a severe threat of harm to others. Both the murderer at the door and the individual in the theater forfeit their entitlement against interference because they breach the standards of co-legislation with other rational agents by assaulting their autonomy.

I will now repeat what I said above using censorship as a stand in for deception. There is not a practical contradiction when censorship is allowed in defense of the autonomy of oneself or others. The foundational value of communication is that it promotes autonomy and allows individuals to realize their own projects and ends. An agent will only practically will a system of effective communication insofar as it contributes to the preservation of a justified autonomous sphere of action for individuals. An agent commits no practical contradiction by censoring in defense of the autonomy of oneself or others because the agent is willing the universally coherent means in the world to respect their autonomy. Such means will largely demand honest communication but will admit of certain principled exceptions. Thus, defending the autonomy of others is a principled basis by which an individual can either lie or censor.

Conclusion

I began the chapter by discussing the fact that humans are agents and this imposes a structure of obligations on our reasoning. As reflective creatures, we are

forced to structure our choices and values in a way that makes our lives intelligible. Reasoning is the process by which we make our lives intelligible so we have obligations to be rational. Therefore, any valid moral theory needs to have a rational structure. I then considered what type of rational structure is necessary for any valid moral theory to have. As the structure of reasoning is accessible to all rational agents, a rational principle needs to be universally valid for all rational agents. This is the principle of universalizability which holds that rational principles must be universally and conclusively reached by all equally rational agents who impartially weigh the considerations at issue. The second principle of rationality is the principle of coherence which states that all rational principles must be coherent and cannot contradict any other rational principles. Therefore, any valid moral principle must be universally coherent.

The argument above is a formal framework of rationality. It claims that all valid moral theories must be universally coherent in order for a rational agent to act intelligibly with respect to it. A weaker claim this chapter presents is that all moral theories must abide by certain standards of rational consistency. This chapter also went beyond simply formal considerations and also defended a more substantive view of lying and censorship. Many different moral theories can pass the formal framework outlined above. However, I argue that these theories are substantively lacking if they do not allow agents to defend their autonomy and the autonomy of others from unjustified interference. I rejected several theories that fail to establish this principle and I ultimately defend a theory that has a principled basis for interfering with others to the extent they violate the autonomy of rational agents.

I claim that every rational agent should be practically committed to an effective structure of communication. An agent commits a practical contradiction if they simultaneously will a world with effective communication structures and also engage in behavior that subverts the value of communication structures. Communication is crucial for realizing many important moral values and so interfering with the communication of others is a *prima facie* wrong. However, there are certain cases where we think strategic communication can be used to interfere with the communication of others. Most people believe it is permissible to lie to the murderer at the door and to censor the speaker who is going to intentionally and falsely yell fire in a crowded theater. The paradox of communication is the need for a moral theory to have a universally coherent position that explains why it is wrong in most cases to interfere with communication but justified to interfere with the murderer at the door and the theater speaker without committing a practical contradiction. I outlined several moral theories which try and explain this paradox.

I argue that the liberal argument for free expression is the moral theory that best solves the paradox of communication. It explains why censorship is usually wrong because it violates the autonomy of others but also provides a principled basis for censorship in both cases we considered. This liberal model is equipped to permit deception in the murderer and Nazi at the door cases. Moreover, its appeal to autonomy through and through makes it a viable rational framework because it rests on a universally coherent principle. Autonomy is the core value that is defended on a principled basis universally for each rational agent. As autonomy is the controlling value, appealing to it avoids a rational contradiction.

The account of the liberal censor that I have outlined is concerned with censoring only in the name of preserving individual autonomy. Certain forms of speech are an unjustified interference against the autonomy of others and can be permissibly censored like in the theater case. I acknowledge that this standard is very vague at the moment. While many might support the liberal model I have outlined in the abstract, there are serious disagreements about what constitutes an unjustified interference against the autonomy of others that permits censorship. In the next chapter I elaborate on these different theories of individual entitlements with respect to the liberal theory and argue that many attempts at censorship under the guise of the liberal model are unjustified. Many attempts to censor in the name of liberalism are contradictory and censoring on that basis would be an act of hypocrisy. The work of the third chapter will be to pare down the liberal principle of censorship into a more unified and coherent form.

Chapter 2: The Hypocritical Censors

Introduction

In the previous chapter I laid the groundwork for censorship that is justified in order to prevent the unjustified infringement of autonomy against oneself or others. I will refer to this moral framework of censorship as the Liberal Censor. In Chapter 3 I will offer a more detailed outline and defense of the Liberal Censor substantive thesis and argue that it is the only justified defense of censorship. However, there are many other arguments for various kinds of censorship. This chapter will consider four of the more prominent censorship arguments. These censorship arguments attempt to be liberal because they want to preserve a government allowing a wide range of expression and one that broadly protects the expression of individuals. They propose limited forms of censorship in the name of what they perceive to be liberal values. I will argue that all of these arguments are hypocritical because they violate the formal rational framework outlined in Chapter 1. Each of these theories censor some speech but fail to censor speech that can be principally censored under the same principle. This position is untenable because if the principle offered by the hypocritical censors is universalized then the hypocritical censor commits a practical contradiction in the type of censorship they will.

Hypocritical censors attempt to will a principle of censorship that cannot be universalized. Hypocrisy is wrong because it fails to treat rational agents with equal moral standing by holding agents to different speech standards that are rationally incoherent. I will argue that in order to avoid hypocrisy these types of censors need to commit themselves to universalizing the principles of censorship they advance. If they do so then they will move from being a hypocritical censor to an ideal censor which

adheres to the formal standards of rationality. Nonetheless, these ideal censors, while formally valid, are not equal substantively. I will argue in Chapter 3 that all the ideal censors except the Liberal censor suffer substantive defects. The Liberal Censor is the acceptable censor formally and substantively.⁴⁵

Many people subscribe to one of the four censorship strategies considered below and accept the permissibility of censorship according to community standards, offense, hate speech, or promoting democratic discourse. I will point out that the principles they attempt to censor speech if universalized would be illiberal. One could, of course, find ways to censor the kinds of speech outlined in this chapter but that would require committing to much more authoritarian and implausible moral principles that go beyond what most want to commit to. The censors in this chapter are hypocritical if they try maintain themselves as liberal censors because they are carving out exceptions for themselves without applying the standards universally and in all cases in which they apply. This is morally wrong because it deprives the people who are victims of this kind of censorship of due process. T.M. Scanlon explains that due process “aims to provide some assurance of nonarbitrariness by requiring those who exercise authority to justify their intended actions.”⁴⁶ If moral principles are not consistently applied then this violates rational standards and remove the justificatory power officials have to censor.

The hypocritical censors act wrongly because they deprive others of due process and subject them to the force of arbitrary authority. Ultimately there is a disconnect between the principles of actions and the character the “liberal” censors impute to

⁴⁵ The third chapter will also make arguments for why the Liberal Censor is superior in terms of ideal and non-ideal moral theory as well.

⁴⁶ T.M. Scanlon, *The Difficulty of Tolerance: Essays in Political Philosophy* (Cambridge University Press, 2003), 44.

themselves. Everyone who rejects the authority for officials to use force in an arbitrary manner should reject hypocritical censorship and demand due process for those who are censored, a demand that censorship is justified according to consistently applied rational standards. This chapter will outline four hypocritical censors who fail to consistently apply their censorship standards because they do not want to act illiberally. I compare these hypocritical censors on the left to the consistent censors on the right of the table. The censors on the left have a corresponding consistent censors on the right one would need to commit to in order to make the application of their censorship principle extensionally adequate. The consistent censors will be discussed in the third chapter and they anchor their moral justifications in deeper values which ultimately deviate in important ways from the liberal censor model.

The Cultural Relativism Censor	Majority Rules Democracy
The Humanitarian Censor	The Utilitarian Censor
The Hate Speech Censor	The Radical Egalitarian Censor
The Democratic Censor	The Maximal Autonomy Censor

Section I: The Cultural Relativist Censor

The first hypocritical censor that I will consider is the cultural relativism censor. This censor believes there is not an objective character to morality and that our practices of speech censorship should be evaluated culture to culture rather than according to universal standards. In the first chapter I argued that we should think of morality objectively and focus on universally coherent standards for assessing the morality of our conduct. The cultural relativist censor pushes back against these

foundational principles. The cultural relativist avoids representing the conduct of people from other cultures as right or wrong because they believe they lack an independent standard which allows for such a characterization. Lacking a universal standard of judgement, the cultural relativist argues that our cultural practices are not special but instead “merely one among many.”⁴⁷

The speech relativist sees differences in approaches to speech regulation in different cultures and claims that we cannot find objective features that decide which approach is objectively superior amidst these differences. For example, the speech relativist might point to the very different cultural approaches to free expression between the United States and Europe. The United States sets very strenuous constraints on when the government can limit free expression whereas most European governments have speech regulations that are more strongly balanced against other social considerations.⁴⁸ At this point, the relativist is inclined to throw their hands in the air and say something along the lines of “America has their speech standards and Europe has theirs” and state that we cannot decide morally between these approaches. These differences in what societies consider to be moral behavior lead cultural relativists directly to the conclusion that there are no objective and universal answers to moral questions. Consequently, cultural relativism strips morality of its objective content that transcend particular human societies. The applicable censorship principle can be expressed as follows:

⁴⁷ James Rachels, “The Challenge of Cultural Relativism,” in *Exploring Philosophy: An Introductory Anthology*, ed. S.M. Cahn, 4th ed. (Oxford University Press, 2009), 3.

⁴⁸ Ivan Hare and James Weinstein, *Extreme Speech and Democracy*, 1 edition (Oxford ; New York: Oxford University Press, 2009), 2.

Cultural Relativist Censorship Principle:

Individuals have no authority or ability to declare which cultural practices are more right or moral than others. Censorship standards should be determined uniquely by each human cultural community.

The idea of cultural relativism leans heavily on the idea of authority and that we are not able to establish that other cultures are objectively wrong in their approach to morality. Accepting the argument for cultural relativism would have serious implications that would erode authority from our concept of morality. If this occurs then morality would be devoid of value. If morality is relative then there are no standards for criticizing the cultural practices of others because there is no independent standard of morality to appeal to. If a culture engages in slavery then we have no moral ammunition to judge or criticize them because the moral views of that society are completely derived from that particular culture. I want to distinguish between two types of cultural relativists: strong relativists and weak relativists. Strong relativists argue that humans have moral obligations to follow the morality of their particular culture. This standard does not work. For example, if you were born into a culture that deemed slavery as a morally essential practice then the strong relativist argument would conclude that you have a *moral obligation* to practice slavery.⁴⁹

A cultural relativist wanting to avoid this conclusion could take refuge in the weak cultural relativist argument. In contrast to strong relativists, a weak relativist would say that morality is the product of human societies and that you as an individual have no moral obligation to comply with your society's cultural dictates about morality. This approach is problematic for two key reasons. First, accepting this

⁴⁹ I do not think any reader will accept this conclusion so I will not discuss strong cultural relativism further.

principle means you as an individual would lack the moral resources necessary to say that your society's practice of slavery is immoral. All you would be able to say is that *you* disagree with what your society is doing but you would not be able to appeal to an objective moral standard that says that slavery is wrong and was always wrong independent of culture. A second issue is that the weak relativist approach fails to effectively ground moral obligations. If there is no objective moral obligation to act in certain ways and there is no societal moral obligation to act in certain ways then there is no independent source of moral obligations outside of individual preference. The view suggests that the only source of binding moral obligation is what the individual places on herself or himself. We do not normally think of moral obligations as being the source of "whatever I determine to be so is so" and typically we think of morality as being independent of ourselves. Moreover, if moral obligations reside with the individual then an individual would lack a moral obligation to abstain from murdering or enslaving others if they do not think that such moral obligations apply to them. If an individual chooses to not feel obligated to be moral then there is nothing independent of them to hold them morally accountable in the weak relativist view. If these conclusions are to be avoided then weak cultural relativism should not be accepted.

For these reasons I think that both strong and weak cultural relativism arguments should be rejected. If there is a character of morality then it needs to be objective and have force independent of any particular individual and culture. We accept such independent standards with science and mathematics. There is no such thing as Indian chemistry or Chinese mathematics. There is only science or mathematics which holds force beyond any individual or culture. Analogously, we should believe there is no such thing as Indian or Chinese morality. Without saying

anything yet about what the moral standards are, we should think there is an objective character to morality and that morality would have universal authority and force across all rational agents and cultures.

The speech relativist might pivot at this point and propose a universal standard of moral authority that would still preserve the concept of speech relativism between different societies. This principle is anchored in a majoritarian conception of democracy that says that a society has the sovereign authority to dictate laws that regulate its social order. I will call this the principle of democratic relativism. The fundamental premise of this argument is that societies have the right to determine the standards which govern them through their majorities. If the majority of Europeans and the majority of Americans in free democratic elections produce a different set of speech laws then both those laws should be acceptable to the democratic speech relativist. The view is that a society gets to require or prohibit certain forms of behavior according to the will of the majority. This relativist position tries to strike a liberal chord. To the speech relativist, it is liberal to respect the self-determined wishes of a society and it smacks of imperialism and arrogance to tell other democratic societies that their established customs and social regulations are objectively wrong.

To avoid the objection I raised to cultural relativism above, the speech relativist must advance a moral principle that is universal and still maintains respect for different speech standards of cultures. The principle that is advanced is that the majority of society is justified in applying laws to shape the social construction of society as a whole. What behaviors that violate a society's cultural code can be adjudicated through democratic elections, referendums, and expression of the majority of the people's will in government. This may involve requiring minorities to

conform to certain modes of behavior by force. The pithy title of this form of democracy is “Majority Rules.” If this kind of democratic principle is advanced then it would create a speech relativism that would ultimately be anchored by a universal standard that all societies have the authority to use coercion to enforce compliance with what is determined by the society’s common moral code. Majorities of different societies have different preferences when it comes to speech they want expressed. American voting majorities would create different speech laws than Europe and create two different legal spheres of expression. Thus, this speech relativist would say that any speech regulation would be adequate if it is produced by a democratic majority in a society.

This argument for democratic speech relativism should be rejected. The reason is that accepting this principle is incredibly shallow and would be a terrible principle to logically extend morally. The principle allows the use of coercion and the regulation of behavior simply because a majority of the people believe that this should be done. This is incredibly problematic because no standards are imposed on who gets to cast a ballot that regulates coercion in a society. A person could be a “village idiot” and still get a say in this theory about the use of force against others who do not wear, say, or think how they like. I use the term “village idiot” because nothing more is required of these voters by the democratic relativist. The principle of majority rules does not say that the majority has to cast ballots constrained by principle. Rather, this standard only requires that the people voting need to have their breath to stick to windows. People only must be a certain age and have the capacity to push a button. This theory is dangerous because it says that a village idiot, be they real or imagined, has a sliver of authority to coerce others on nothing more than the principle that they say so and desire certain things. Most people would not accept this shallow view of democracy.

A morally adequate view of democracy accepts constraints on how democracies are able to permissibly act and fundamentally limit the ability of majorities to infringe on the rights of minorities. If individuals have rights then those rights offer a bulwark against what capricious majorities can do. Therefore, the possession of individual rights would lead to the conclusion that the use of power by a majority cannot be untrammelled.⁵⁰ If “majority rules” is an absolute principle of government then there would not be any end to what government can authorize in the name of the majority. There would be no such thing as the rule of law in this type of government because the rule of law would collapse in the face of a government that can dramatically and suddenly change the law to the whims of a majority. James Madison notes that democracies require a “safeguard against the tyranny of their own passions” so that they escape “the indelible reproach of decreeing to the same citizens the hemlock on one day and statutes on the next.”⁵¹

We do not accept that majorities can simply trample on the rights of individuals because they are in the majority. It is ridiculous to say that who someone can marry, what they can wear, and what they are able to say can be simply overturned because they were outvoted. Michael Huemer rightly points out that a vote of a majority of people is inadequate to confer legitimacy to use coercion against others. He imagines a case where you walk into a bar and someone suggests that you pay for everyone’s drinks in the bar. Over your objection, the majority of people vote to have you pay for

⁵⁰ I recognize that I owe a further defense to the claim that individuals have rights. I will pick up this discussion in chapter 3. I leave the claim here that individuals have rights because I believe that most readers believe that individuals have rights that are derived independently from what a majority dictates.

⁵¹ Alexander Hamilton, James Madison, and John Jay, *The Federalist Papers*, ed. Clinton Rossiter (New American Library, 2003), 383.

the drinks.⁵² Of course, you have no obligation to pay for everyone's drinks because the bare principle of a majority rules is ridiculous. A majority, just because it is the majority, cannot use force to regulate bar tabs, social behavior, and speech. Therefore, we should reject democratic relativism because its underlying principle is not a plausible one.

In this section I considered three variants of the argument for relativism. These relativists attack the whole project of finding moral principles for speech regulation from the outset. They think that societies are in a unique position to dictate the kinds of speech that are acceptable and use coercion to enforce their particular cultural view. All of these variants are incorrect. Strong relativism states that individuals have a moral obligation to follow the policies of their own culture. This proposition would assert that individuals have the moral obligation to abide with laws like slavery if that is what their culture prescribes. Weak relativism abandons the obligation argument but contends that there are no objective moral standards allowing one to objectively criticize their culture or the cultures of others. Accepting weak relativism would mean retreating from morality in terms of objective standards and this would shatter our ability to make effective moral criticisms and judgements. Finally, democratic relativism is the view that majorities in societies have the right to prescribe laws and use coercion to enforce them. This majority rules conception of democracy would allow speech regulation if the majority of society supports it. However, this principle should be rejected because a majority by itself has no inherent moral legitimacy. Accepting this view of democracy would require forsaking the concept of individual rights

⁵² Michael Huemer, *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey* (Palgrave Macmillan, 2013), 75.

because the majority rules view of democracy promotes the untrammelled use of power of majorities in a democracy.

I will later discuss a richer account of democracy that accepts fundamental constraints on its power. My aim in this section is to dismantle the shallow and feeble account of democracy that grants authority to majorities because it is a majority. This Majority Rules conception of democracy cannot stand according to a plausible universally coherent principle. If this principle cannot stand then we would not be able to say that “Europe has their speech standards and we have ours” because the majority of Europeans elected for a particular approach. Differences in standards relating to speech relativism cannot withstand moral scrutiny because the moral principle they rest upon is hollow. With the relativism arguments being deficient we should search for standards of speech regulation that are grounded in universally coherent principles that can be logically extended objectively across all human societies and across all persons. The remainder of chapter 2 will consider three further arguments that fall short of this standard.

Section II: The Humanitarian Censor

Most discussions of the justification for censorship usually involve questions of harm. John Stuart Mill famously defended the right of free expression in society and argued in favor of the benefits of a free society where the state’s authority to curtail speech is limited.⁵³ Thus, Mill claimed that protecting free expression would benefit society more on balance than limits on such expression. However, Mill also claimed

⁵³ Mill, *On Liberty*.

that free expression could not be unlimited and allowed for the state to limit speech which “without justifiable cause, do harm to others.”⁵⁴ He goes on to say that “The liberty of the individual must be thus far limited; he must not make himself a nuisance to other people.”⁵⁵ These statements help to form the basis of Mill’s Harm Principle where the state cannot liberty for a person’s own good but that they are able to limit liberty when actions pose an unjustifiable risk of harm to others. Such a distinction is not very helpful to us unless the types of harm that can be prohibited are specified. This is because one’s view of harm can either be expansive or narrow. There are at least two conceptually distinct forms of harm that speech can cause: rights based harms and welfare based harms. I will argue that Mill’s Harm Principle can be coherently extended for rights based harms but should not be coherently extended for welfare based harms. The person that wishes to coherently extend the Harm Principle for welfare based harms becomes the Utilitarian Censor which I will discuss in more detail in Chapter 3.

Mill is correct that all speech cannot be tolerated because some kinds of speech cause unjustifiable harms. Jeffrey Howard argues that “the moral right to freedom of expression, properly interpreted, does not protect speech that incites clear violations of others’ moral rights.”⁵⁶ Unjustifiable harms are harms that violate the rights of individuals and certain kinds of speech can violate rights. As Howard notes, speech restrictions are acceptable when they prevent “incontrovertible” rights violations that infringe on the bodily autonomy of individuals.⁵⁷ Chapter 1 highlighted the Fire Case

⁵⁴ Mill, 121.

⁵⁵ Mill, 121.

⁵⁶ Jeffrey W. Howard, “Dangerous Speech,” *Philosophy & Public Affairs* 47, no. 2 (2019): 210, <https://doi.org/10.1111/papa.12145>.

⁵⁷ Howard, 217.

where a speaker intentionally and falsely yells fire in a crowded theater to start a panic and harm others. This is a case where individual rights are “incontrovertibly” violated and the bodily autonomy of individuals are directly threatened. Now let’s consider another case.

The Bullhorn Case:

Frank believes that taxation is theft and that the government violates his rights when it taxes him. He wishes to express his displeasure of government taxation policies so he finds the address of an IRS agent named Sue. At 2:00 AM he goes outside the IRS agent’s home and uses a bullhorn to loudly complain about taxation policies which wakes up Sue and prevents her family from sleeping.

I think it is clear that Sue’s rights are violated by Frank’s speech based on the context in which he spoke. By using a bullhorn to disrupt the agent’s sleep and to loudly speak at night outside the agent’s home she is not able to reasonably avoid the speech. It is not the substance of Frank’s comments which violate the agent’s rights but rather the *manner* in which he presents them. Frank is conducting a kind of assault by directing sound waves in an unjustifiably disruptive manner. Sue is entitled as an autonomous agent to a certain sphere of control where she can exercise her bodily autonomy. Frank infringes on this sphere of control by projecting harmful noise at night where Sue should have the right to sleep in peace in certain areas like her home that within her sphere of control. Nonetheless, a person’s sphere of control will vary with the particular environment they are in. For instance, if Frank is yelling in the middle of the day in a public park about how taxation is theft and Sue hears it while passing through a public park her rights are not violated. Likewise, Sue does not have her rights violated if she receives a letter from Frank accusing her of violating the

rights of individuals and stealing from people. Her rights are not violated in these latter cases because the sphere of control in which the agent operates has changed. Sue lacks authority to control messaging in a public park and as a public official she does not have the right to exclude public commentary on her actions. In chapter 3 I will elaborate more on a theory of rights and when individuals have claims against others for speech. For the time being I will consider the argument that individuals can also be censored for speech that causes general welfare based harms.

I titled this section the Humanitarian Censor because it is drawn from Jonathan Rauch's discussion of individuals who are concerned with humanitarian or welfare based harms from the speech of others.⁵⁸ The Humanitarian Censor takes a more expansive view of harm and thinks that speech that undermines the welfare of others can be permissibly regulated. Speech can also cause offense that affects the sensibilities of others. One may take the approach that the state has the right to limit any speech that can provide a setback to the welfare interests of any other individual. Yet, this principle would be absurd if it were to be extended because an individual's welfare interests are dependent on factors that do not provide a reasonable justification for censorship.

To see why restricting speech for welfare based harms is extensionally inadequate let us return to the Bullhorn Case. The IRS agent Sue has her welfare linked at least in part by what other people think of her. Her welfare drops when others speak poorly of her and when her reputation is harmed. The extent to which her welfare changes is moderated by her personal feelings that are subjective to her as a person. Put more bluntly, some people have thick skins while others have thin

⁵⁸ Jonathan Rauch, *Kindly Inquisitors: The New Attacks on Free Thought*, 2nd ed. (The University of Chicago Press, 2013), 111.

skins. Imagine that Sue has a fellow IRS agent named Sarah who has incredibly thick skin. Sarah can sustain blistering criticism and shrug the words of others off without any effect on her welfare. In contrast, Sue has very thin skin. She is like an eggshell and will shatter in the face of even slight criticism. While Sue is on her way to work as an IRS agent she encounters Frank in a public park who is holding a sign that says that those who work for the IRS are thieves and are violating the rights of individuals. Upon seeing this sign Sue shatters. The allegation that she is a thief devastates her welfare and makes her incredibly upset. We can imagine that she enters work and is unable to focus, her relationship with her colleagues suffers, and her family no longer brightens her day as much as they used to. Sue has clearly suffered severe harms to her welfare and well-being as a result of Frank's speech. If the government has the moral license to censor speech that undermines the welfare of others then it would be clearly justified in censoring Frank's speech. Frank derives a small amount of welfare gain from holding up the sign but its effect is absolutely devastating to Sue.

The harm to Sue's welfare alone does not provide an adequate reason to censor Frank's speech and to prevent him from expressing his opinion. If one were to accept the moral principle that the state should censor any speech that causes significant welfare based harms then one would be morally committed to prevent Frank from speaking in the public park. To adopt this approach, one would have to come up with a consistent ground for censoring welfare based harms across persons. The Utilitarian Censor that I will discuss in chapter 3 does just this by using utility as the framework for comparing the benefits and costs of actions. This type of ideal censor claims that welfare based harms when it comes to utility are a permissible ground for censorship. I will set aside this argument for now until the third chapter.

Having initially rejected the principle of censoring speech on the basis of welfare based harms broadly construed, I now turn to the argument that offensive speech can serve as a grounds for censorship. The Humanitarian Censor is sympathetic to censoring certain kinds of offensive speech because offense can negatively impact the welfare of those subjected to the speech. Joel Feinberg argues that actions that cause certain forms of offense can be permissibly regulated by the state even if they are not harmful. This occurs because what is offensive and what is harmful are not necessarily related.⁵⁹ To support this position Feinberg outlines a framework for evaluating when offensive conduct can be prohibited.

The Offense Principle:

“It is always a good reason in support of a proposed criminal prohibition that it would probably be an effective way of preventing serious offense (as opposed to injury or harm) to persons other than the actor, and that it is probably a necessary means to that end.”⁶⁰

Feinberg clarifies that the Offense Principle relates to wrongful offense where offense is a disliked state caused by another that is worthy of resentment.⁶¹ Certain speech acts cause offense because there are a multitude of expressions which cause disliked states that individuals believe are worthy of resentment. However, the Offense Principle has to be moderated and cannot assert that *any* offense to a particular individual of a certain magnitude can be prohibited. There are several reasons for this and they are based in the fact that offensiveness is a subjective emotional experience. First, offensiveness varies between agents due to background experiences and what

⁵⁹ Joel Feinberg, *The Moral Limits of the Criminal Law: Offense to Others* (Oxford University Press, 1988), 2.

⁶⁰ Feinberg, 1.

⁶¹ Feinberg, 2.

agents find to be symbolically important. For example, a person saying that “Jesus Christ was an evil cult leader” could deeply offend a devoted Christian while making an atheist or even another Christian smile. Second, the magnitude of offense varies from person to person. Recall the case of Sue and Sarah above. Sue has an eggshell personality when it comes to offensive statements about the IRS whereas Sarah has a very thick skin to such slights. A negative statement about the IRS could offend both Sue and Sarah but Sue would be much more seriously affected. Finally, offense even varies in affect and magnitude for individual agent’s themselves over time. When I would visit the zoo as a young child I loved visiting the gorilla exhibit and was particularly enthralled when the gorillas would throw up in their hands and then eat their vomit. Now I find that I am revolted watching that process as an adult. Individuals can change their life outlooks over time. A young person may be an atheist at a young age and not be offended by criticism of Christianity but then become a devoted Christian later in life and be deeply offended by similar statements. Thus, the severity of offensiveness is moderated by an individual’s background, their particular resiliency against what can be offensive, and the changes an individual experiences in these dimensions over time.

The features of offense that change from person to person poses a serious challenge for creating a moral principle for regulating free expression. Chapter 1 established that rational moral principles must be universally coherent. But offense is extraordinarily contingent and speech cannot have a universally offensive effect. As Justice John Harlan once said “One man’s vulgarity is another’s lyric.”⁶² One could resolve this difficulty by saying that severely offensive speech of any form can be

⁶² “Cohen v. California, 403 U.S. 15 (1971),” Justia Law, accessed February 28, 2021, <https://supreme.justia.com/cases/federal/us/403/15/>.

prohibited by the state. I will call this the broad offense principle. This principle does not work because the principle, unsurprisingly, would be too broad. Let's return to the IRS case. Imagine that when Frank holds up his sign "IRS agents are thieves" in a public park in the proximity of Sue he causes severe offense. When he does so around Sarah there is minimal offense. If Sue walks to the park on Tuesdays and Thursdays then a broadly construed offense principle would say that the state can coerce Frank when he displays the sign on Tuesdays and Thursdays but not on days when Sue is walking through the park. This is a ridiculous standard because Frank would have no reasonable way of ascertaining when his speech would cause severe offense and because one person's offense should not dictate when someone can engage in basic freedoms to express oneself.

Another problem of accepting this broad offense principle is that statements like "women should have equal rights" that deeply offend a committed misogynist would fall under the state's authority to censor because it would involve severely offensive speech. Additionally, certain expressions in a public park like a gay or interracial couple expressing affection could severely distress a homophobe or racist. A moral principle that lets the homophobe or racist's distress in this case provide a reason for censorship is no valid principle at all. The state cannot regulate speech on the basis of welfare harms from offense because the regulation would contingently vary from person to person and limit expressions that we believe individuals are entitled to express.

Based on these concerns, Feinberg offers principles that mediate the Offense Principle to offer a more coherent principle that does not invite "wholesale and

intuitively unwarranted legal interference.”⁶³ Feinberg proposes that offensive conduct should be mediated by the following factors:⁶⁴

1. Whether the offensiveness is of a serious magnitude.
2. Whether those who are offended by the conduct can reasonably avoid being offended.
3. Whether or not those offended willingly assume the risks of being offended.
4. Whether or not the person being offended is a standard observer.

Feinberg thinks that offensive conduct can be prohibited when it creates offense of a serious magnitude that a standard observer who has not willingly assumed the risk of being offended cannot reasonably avoid being offended. I will call this the Modified Offense Principle (MOP). The MOP offers more useful resources for avoiding the unwelcome conclusions reached through the broad offense principle. The MOP would not allow the state to censor Frank’s speech on the basis of offense because Sue is not a standard observer and she is able to reasonably avoid the speech by not traveling through the public park. A lot of offensive speech that occurs in specific areas like adult theaters, books, or internet sites would also be protected from the MOP because standard observers would be able to reasonably avoid being exposed to such speech.

Feinberg uses the MOP to justify regulations of certain kinds of severely offensive speech like the Nazis marching and displaying anti-semitic symbols in the predominantly Jewish neighborhood of Skokie. Feinberg thinks that the offense in Skokie justifies regulation even if no one is harmed and no rights are violated because the conduct passes satisfies the MOP and there are no sufficient social interests in

⁶³ Feinberg, *The Moral Limits of the Criminal Law: Offense to Others*, 27.

⁶⁴ Feinberg, 26–27.

hearing the speech to “balance” it against the offense caused. The MOP is a standard that attempts to maintain a liberal state where speech is broadly protected. At the same time, the MOP functions to exclude certain severely offensive speech from public discourse.

I disagree with Feinberg’s modified offense principle and I do not believe that it can provide an adequate moral principle for censoring speech. I argue that offense is *never* a justifying reason for social regulation of individual conduct. Offense may be coupled with conduct that can be prohibited but never provides an independent reason for legal prohibition. For instance, a stranger may slap me and I may be offended that he did so. The stranger may be criminally punished for slapping me but he may not be punished or his punishment exacerbated because I took offense. The reason why the MOP should not be accepted is that it would generate justifications for state authority that are simply unpalatable. For example, consider the following case:

The Gruesome Man

Rick has lived an unfortunate life. When he was younger his house caught fire killing his family. While Rick escaped, he suffered horrendous and permanent burn injuries. His appearance is now gruesome and standard observers look at him they cannot help but to shudder and feeling a sense of revulsion. In short, Rick’s appearance is offensive on sight because he causes others to suffer unpleasant feelings when looking on him that lead them to resent Rick’s presence. His disfigurements cause more offense and revulsion than even the public display of the most hard core pornography. Rick enjoys taking walks in the public park where people see him and are severely offended by his presence. One day a police officer comes up to Rick and gives him a criminal citation after having warned him one time before against walking in the park. By publicly

*displaying his appearance Rick is causing offense to others under Feinberg's modified offense principle and is now liable to be coerced by the state.*⁶⁵

I am assuming that the reader will find it unconscionable that Rick should be criminally liable in principle for walking in a public park. Nonetheless, Rick's actions do fall under the MOP. The offense he causes is of a serious magnitude to a standard observer, others find it hard to reasonably avoid seeing him in a public park, and the risk of being offended by Rick's appearance is not willingly assumed.⁶⁶ I think the Gruesome Man case points to a reason why the MOP should be rejected. Even if it causes offense, Rick's appearance offending others is not a reason at all for criminal prosecution.

Another key problem with Feinberg's MOP is the mediating principle that says that severe offense must be experienced by a "standard observer." This is because what is a standard observer depends on how wide one casts a net over an area to survey the observers to determine what is standard. The standard observer would dramatically change if the area surveyed was a tiny religious town in Alabama, or the entire state of Alabama, or the entire United States, or the global population. Which of these "standard" observers shall we select? Feinberg does not propose an adequate mechanism for determining what a standard observer would be and the area one selects can lead one to opposite conclusions under the MOP. Making a joke at the expense of the Prophet Muhammad may not be severely offensive if your standard observer is in the United States but would be extremely offensive if one's sample size for standard observers come from Iran. The larger one goes in terms of the area

⁶⁵ I am grateful to Nancy Schaubert for the inspiration for a version of this case.

⁶⁶ If one does not think that all these tenets are met by the example then one can always modify the example to the reader's preference like moving Rick's location to the only neighborhood supermarket where a case for reasonable avoidance is much harder to demonstrate.

sampled makes it more difficult to determine what is standard. After all, the vast diversity across the globe and even in the United States alone means that there really is no “standard” observer that one can mechanistically average to make determinations of criminal liability. Making the sample size narrower creates a whole new set of problems because you are now empowering a small group with the power to criminalize subjectively determined offensive conduct. For instance, if your sample size is Iran or a small religious town in Alabama then the sight of a gay couple holding hands would likely be extremely offensive to “standard” observers there and would provide a reason for criminal prohibition of the offensive conduct. Thus, Feinberg’s MOP should be rejected because of its inability to formulate a standard observer.

Feinberg’s MOP should also be rejected because accepting it leads to a dilemma with two unpalatable options. The MOP can be thought of as either a formal or substantive constraint on the censor. A formal constraint says that so long as the four conditions of the MOP are met then that provides a reason for censorship. A formal constraint is inadequate because it would not be able to rule out certain forms of expression to be suppressed as offensive. If Frank walking in a public park or a gay couple holding hands causes a sufficient kind of distress of a certain kind then it cannot be formally excluded from the MOP. Accepting the principle of limiting offensive conduct as a formal constraint on the censor means accepting a chilling infringement on liberty. As Orwell once said “If liberty means anything at all, it means the right to tell people what they do not want to hear.”⁶⁷ There are countless cases of individuals being offended by the truth. Justice Louis Brandeis emphasized that “the

⁶⁷ “Freedom of Speech Quotes,” Goodreads, accessed February 28, 2021, <https://www.goodreads.com/quotes/tag/freedom-of-speech>.

function of speech is to free men from the bondage of irrational fears.”⁶⁸ Offense itself can embody the subjective emotions of people that reflect their irrational fears. Prohibiting offense would then pose a direct threat to the functional value of speech itself. Moreover, even if speech is wrong there is still value in its autonomous exercise and wrong to limit autonomy on the grounds of offense alone. The wrong of censoring on the basis of offense is encapsulated in the case of the United Kingdom where authorities criminally cited a 15 year old boy for peacefully holding the sign that read “Scientology is not a religion, it is a dangerous cult.”⁶⁹ There is no doubt the boy’s speech deeply offended some. The authorities justified their actions by saying that the community had a “right” not to be “alarmed” or “distressed” by others.⁷⁰ The distress the sign causes is no reason for criminal prohibition. Even if the boy was wrong and Scientology was not a cult or all of its tenets were objectively true that would still not be a good reason to censor the boy. There are some basic liberties, like the boy holding his sign, a gay couple holding hands, or Rick walking in the park, that should not yield to the offense of others. Offense based concerns are absolutely silenced by certain entitlements to expression.

One could avoid the challenges of making the MOP a formal constraint by making it a substantive constraint. A substantive constraint would say that some types of offense provide no reason at all for censoring speech. For instance, a substantive constraint could say that the fact that a racist, sexist, or homophobe are offended do not provide reasons for censoring speech. Additionally, this position could declare that certain expressions, like Rick walking in the park, are categorically

⁶⁸ “Whitney v. California, 274 U.S. 357 (1927),” Justia Law, accessed December 27, 2020, <https://supreme.justia.com/cases/federal/us/274/357/>.

⁶⁹ Hare and Weinstein, *Extreme Speech and Democracy*, 40.

⁷⁰ Hare and Weinstein, 40.

permissible and will not yield to offense based considerations. The problem with this substantive approach is that it violates the moral framework of universal coherence. It diminishes certain kinds of offense and elevates certain kinds of expression. In effect, a substantive MOP would require importing a whole new moral principle which would say when it is acceptable to dismiss certain kinds of offense, like that suffered by the racist, and when certain kinds of expression, like walking in the park as you are, must be tolerated by society.

If one imports a new moral principle then that new moral principle does all of the moral work in the case and not offense. For example, imagine that it is important to several women that they get to jog each morning topless through the park. The sight of the topless women causes several birdwatchers in the park to be deeply offended. The women respond that, like Frank, they are just inhabiting the park “as they are” and that it is important to them to enjoy the park as they are. Offense does no substantive work in deciding the claims of Frank and the topless women in these cases. What matters is whether these individuals are entitled to be in the park as they are. If offense is said to form a basis for determining entitlements then we are back to offense functioning as a formal constraint with all of its accompanying problems discussed above.

We do not think that Rick’s offensive appearance provides a reason for him to be removed from the park but that other considerations outweigh. Rather, we simply think Rick has the right to be in the park. Similarly, we do not “weigh” the sexist’s offense at being exposed to feminist speech and then “balance out” the considerations of the feminist to let her speak. Offense is ultimately irrelevant in how we construct

considerations of censorship. Thus, we should never think of offense as a justifying reason in itself to censor speech.

Section III: The Hate Speech Censor

In the previous section I raised the case of Nazis marching in Skokie as a case where one may be inclined to censor the speech of those individuals according to the Offense Principle. However, we rejected the Offense Principle as a principle because it was not an extensionally adequate moral principle. The Hate Speech Censor is worried about speech like the Nazis at Skokie being expressed. Andrew Altman points out that “even when it involves no direct threat of violence, hate speech can cause abiding feelings of fear, anxiety, and insecurity in those at whom it is targeted.”⁷¹ The Hate Speech Censor seeks to limit hate speech which targets the dignity of certain groups while still preserving a social structure that is broadly tolerant of dissent and the open exchange of ideas. Censoring on the basis of offense requires positional judgements because who is offended can depend on the subjective position of those offended. However, censoring for human dignity has an advantage because human dignity does not require a positional perspective and can be universally upheld across all rational agents. Thus, the hate speech censor attempts to advance a universally coherent moral principle that allows censorship in defense of human dignity. I will argue that the Hate Speech Censor is unable to preserve this balance between censorship for dignity and liberal tolerance because the moral principles underlying these approaches are fundamentally in tension.

⁷¹ Andrew Altman, “Liberalism and Campus Hate Speech: A Philosophical Examination,” *Ethics* 103, no. 2 (1993): 309.

Jeremy Waldron has prominently argued that censorship of hate speech can be justified within liberal societies. He argues that all people are entitled to dignity which consists of their “social standing, the fundamentals of basic reputation that entitle them to be treated as equals in the ordinary operations of society.”⁷² Waldron thinks that one’s dignity or social standing is a public good based on the assurance that other members of society respect the social standing of others.⁷³ He defines hate speech as publications “which express profound disrespect, hatred, and vilification for the members of minority groups.”⁷⁴ Hate speech has the ability to undermine a society’s public good for assuring equal social standing for all members in a society. Thus, Waldron argues that censorship of hate speech can be justified in order to maintain an egalitarian sense of dignity within a society. This reasoning would support the censorship of the Nazis at Skokie because their speech and waving Nazi flags would profoundly disrespect minority groups and undermine their confidence that their society respects their equal claim to dignity. Waldron’s approach offers a different moral principle that does not seek to censor the Nazis at Skokie for offensive conduct. Rather, he would justify their censorship on the basis of speech threatening the entitlement of people to participate in society with a basic assurance to dignity or social standing.

I disagree with Waldron’s argument on several levels. My first problem with his argument is that a person’s dignity is not based on social assurance. Dignity is something intrinsic and irrevocable in virtue of an agent’s rational status and is a birthright of a person’s humanity. If others can think or speak ill of you and you

⁷² Jeremy Waldron, *The Harm in Hate Speech* (Harvard University Press, 2014), 5.

⁷³ Waldron, 93.

⁷⁴ Waldron, 27.

suddenly lose your dignity then that is a shallow conception of dignity. My conception of human dignity nods back to Kant who thought that our rationality allows us to realize value in the world when we make objects the subject of our rational choice. Christine Korsgaard explains that for Kant “we *confer* value on the objects of our rational choices” and to “choose something is to take it to be worth pursuing; and when choose things because they are important to *us* we are in effect taking *ourselves* to be important.”⁷⁵ This is crucial because our value as agents, the source of our dignity, comes from within us.

The implication of this Kantian argument is that our dignity is not conditional on the social assurance of others but rather on what we choose to value. If we value the hate of others or give it credence then it will assault our personal standing. Alternatively, if we choose to ignore the hatred of others then it will never be able to threaten our dignity as rational agents because we have *chosen* to ignore that vitriol. As Eleanor Roosevelt once eloquently noted, “No one can make you feel inferior without your consent.”⁷⁶ If we choose to accept the substance of hate then it will affect our social standing. However, if we choose to not let hatred interfere with our perception of ourselves then one’s dignity will be impregnable to the speech of others. Hate speech can never inherently or unconditionally assault the dignity of others. The upshot is that censoring hate speech because of its threat to the dignity of others can never be based on a universally coherent moral principle.

Another problem with Waldron’s position is with how he defines the concept of hate speech. Recall that Waldron definition of hate speech involves speech expressing

⁷⁵ Korsgaard, *Creating the Kingdom of Ends*, xi–x.

⁷⁶ Nadine Strossen, *Hate: Why We Should Resist It with Free Speech, Not Censorship* (Oxford University Press, 2020), 171.

“profound disrespect, hatred, and vilification for the members of minority groups.”⁷⁷

The trouble with this definition is that it allows only expressions which target minority groups to qualify as hate speech. Hateful rhetoric directed at majority groups would not fall under the definition of hate speech to be subject to censorship. But the difficulty discerning whether a person is part of a majority or minority runs into the same observer issue that Feinberg encounters when he attempts to define a “standard observer.” Whether a person is part of a majority or minority group is dependent the frame of reference one takes. A Jewish person from Israel is a member of a majority group when one’s reference is the area of Israel and Palestine. However, if your frame of reference is the entire Middle East then that Israeli would be considered a member of a minority group. A Christian living in a town of atheists would be a minority group member if you only consider that town but suddenly becomes a majority group member if you shift your frame of reference to the United States and suddenly back again to a minority if you take a global perspective. Any adequate moral principle has to be universally coherent. However, Waldron’s definition of hate speech violates the principle of universality by proposing a principle that fails to universally apply a moral standard across all persons. The attempt to define what a majority and minority group is wholly dependent on an arbitrary frame of reference.

Even if one could choose a frame of reference for deciding what a majority or minority group was there would still be a problem because it does not make much sense to make what people are allowed to say dependent on a fluctuating and arbitrary amount of people around them. To illustrate this point let us look at the following case:

⁷⁷ Waldron, *The Harm in Hate Speech*, 27.

The Hateful Proselytizer:

Jane is a fervent Jehovah's Witness and has very strong convictions about her faith. She is repulsed by Catholic teachings. Jane thinks that the effectiveness of Catholic teachings will end up being convincing and condemn her neighbors and friends to being outside of the grace of Jesus Christ. She goes into a town that is predominantly full of Jehovah's Witnesses with a very small Catholic minority. Jane puts up very large signs that says things like "Catholics are agents of Satan," "No loathsome Catholics are welcome here," and "Cast away Catholic filth." Several Catholics see the signs and appeal to the government to censor Jane. They argue that Jane has breached hate speech laws because Jane's speech assaults their dignity and is "profoundly disrespectful" and vilifies their minority group. The state agrees and destroys Jane's signs and jails her.⁷⁸

I will argue that the state should not censor Jane's speech and has no plausible rational basis for jailing her. Jane certainly runs afoul of Waldron's definition of hate speech due to the character of her attacks against a minority group. This case shows the challenge of defining what a minority group is. Catholics are a dominant group in many areas while within this imagined town they comprise a very small percentage of the population. Moreover, population dynamics can quickly alter whether or not speech is hate speech. Imagine that a week after Jane's sermon the town experiences a mass of conversions to Catholicism and a large influx of immigrating Catholics. Now Catholics comprise 51% of the town's population. It appears absurd that Jane can be

⁷⁸ This case is loosely based on the Supreme Court case *Cantwell v. Connecticut*. There are several major differences that have been modified for the purposes of this argument. In the real case a Jehovah's Witness was speaking in a predominately Catholic neighborhood. The Jehovah's Witness asked two Catholics to play a recording which accused Catholics of being "instruments of Satan." The Catholics agreed but then pressed charges after hearing the recording for a "breach of the peace" which resulted in a conviction. The United States reversed the conviction. See Hare and Weinstein, *Extreme Speech and Democracy*, 35.

censored one week but not the next when the Catholics make up a slim majority of the population. Thus, the number of members of a particular group appears to provide no independent justification for whether or not speech should be censored. The number of members that are in a group and subjected to speech is morally arbitrary.

What this example helps demonstrate is that there is no compelling standard for restricting hateful rhetoric only against groups that are minority groups. But a grant for the state to restrict *any* hateful rhetoric that demeans the identity of a group is way too broad. For example, consider six examples of speech Hateful Jane could put on her sign:

1. "No loathsome Jews are welcome here"
2. "No loathsome Blacks are welcome here"
3. "No loathsome homosexuals are welcome here"
4. "No loathsome whites are welcome here"
5. "No loathsome Catholics are welcome here"
6. "No loathsome Boston Red Sox fans are welcome here"

My guess is that Waldron formulated his definition of hate speech to allow prohibiting the first three statements but not the latter three and certainly not the sixth statement. The problem is that all the statements demean a personal identity of a group individuals are affiliated with. Furthermore, there is no clear way of classifying which statements are more of a threat to a person's dignity because it is dependent on what aspects of a person's identity they care about. A person normally does not treat having green eyes as a salient part of their identity that should be publicly assured because they do not care about their eye color that much when it comes to conceptualizing their identity.

If identity is based on how individuals conceptualize it then the strength of one's identity varies with the strength of individual conceptualizations. For instance, a particular Boston Red Sox fan may be Jewish but care much more about his identity as a Boston Red Sox fan than his Jewish identity. He may not bat an eye about slights against Judaism but be ready to crush the skull of anyone who would demean his beloved Boston Red Sox. It is very condescending to define what the greatest threats to an individual's dignity are for them because a person's dignity is the quintessential thing that is up to the individual. Any attempt to establish a "hierarchy of dignity" and only restrict the worst assaults on dignity would destroy dignity as a valuable concept. There is no clear basis for saying that human dignity should fall into a hierarchy at all. Even if human dignity were to be hierarchical there is no justified basis for saying that government officials have a better grasp on defining the objective "hierarchy of dignity" than individuals do. A view that would attempt to classify a kind of hierarchy of dignity would be committed to an objective dignity standard. However, adopting such a hierarchy of dignity is not feasible and it would mean that any standard that develops the hierarchy is arbitrary and would not make for a plausible standard to serve as the basis of censorship.

There is not a plausible objective dignity standard. One could adopt a subjective dignity standard and say that threats to dignity are up to the individual and can justify censorship. But such a standard would not be tenable because it would be overly inclusive. Way too many cases of speech would be subject to censorship if individuals get to censor anything they consider to be an unwarranted attack on their identity. Severity of dignity based harm also would not plausibly work because an individual Boston Red Sox fan can suffer a greater dignity based harm with the attack on his sports team than an attack on his Jewish identity. If being a supporter of the

Boston Red Sox is a large part of an individual's subjective identity then attacks on the Red Sox could threaten that person's dignity and justify censorship according to this argument. Accepting a subjective dignity standard could allow the state to censor all six of the sentences Hateful Jane offers above. One could make this choice consistently but would end up sacrificing the moral plausibility and moral legitimacy of their position. A person who allows the state to censor all six of the hateful sentences above because they are all "hate speech" would give up any claim to advocating for a broadly liberal and tolerant society. Therefore, I conclude that there is no morally plausible and consistent basis for censoring hate speech if one also wishes to belong to a broadly tolerant and liberal society.

Even if one were to reject the foregoing arguments and say that there is a justified moral standard for hate speech regulation we still should not grant the government the power to censor hate speech. There are practical objections to consider because of the abuse of government authority when it comes to censorship. Governments usually derive most of their power in service of the majority or powerful interests in society. As David Cole notes "in a democracy, the state acts in the name of majority, not the minority. Why would disadvantaged minorities trust representatives of the majority to decide whose speech should be censored?"⁷⁹ There are plenty of examples of hate speech regulation used explicitly against minority groups as Nadine Strossen has extensively researched. In Rwanda hate speech regulation was imposed but then criticisms of the hate speech legislation were deemed "hate speech" that could have been censored.⁸⁰ The University of Michigan implemented a hate speech code on its campus. Enforcement of the Michigan code was turned against minorities

⁷⁹ Strossen, *Hate: Why We Should Resist It with Free Speech, Not Censorship*, 81.

⁸⁰ Strossen, 85.

with 20 instances of white students charging Black students with racist or discriminatory speech including for one Black student calling a white student “white trash.”⁸¹

Regulating hate speech can also lead to the opposite of its intended effect. Timothy Garton Ash reports that studies in England found that “the logging of racist incidents in British schools had the perverse effect of racialising children’s perceptions of each other.”⁸² Maleiha Malik reviewed hate speech regulation and discovered that “incitement legislation often fails to protect powerless minorities” and that “provisions such as incitement to racial and religious hatred are used more frequently to criminalize the speech of minorities rather than protect them from hate speech.”⁸³

Beyond the immediate risk of legislation being used against minority groups, hate speech regulation should also be avoided because it is not effective. All hate speech regulation limits liberty because it narrows what individuals are able to say. Limiting speech based on individual identities and group affiliation is problematic because the 2015 European Commission Against Racism determined that counter speech is “much more likely” than legal prohibitions “to prove effective in ultimately eradicating hate speech and its potential harmful effects.”⁸⁴ Many European countries criminalize hate speech without seeing corresponding gains in social tolerance. The Human Rights Watch reports that “there is little connection in practice between draconian ‘hate speech’ laws and the lessening of ethnic and racial violence or tension.”⁸⁵ The United Nations High Commissioner for Human Rights states that hate

⁸¹ Strossen, 89.

⁸² Timothy Garton Ash, *Free Speech: Ten Principles for a Connected World* (Yale University Press, 2016), 222.

⁸³ Hare and Weinstein, *Extreme Speech and Democracy*, 105.

⁸⁴ Strossen, *Hate: Why We Should Resist It with Free Speech, Not Censorship*, 24.

⁸⁵ Strossen, 137.

speech regulation appears to not “have made a meaningful contribution to reducing racism or... discriminatory conduct.”⁸⁶ Countries like the United Kingdom, France, and Germany all have explicit laws against hate speech but these laws did not stop rises in hate crimes and anti-immigration sentiment against minority communities. In 2019, the number of hate crimes doubled in England and Wales.⁸⁷ Hate speech laws exist across Europe as homophobic and transphobic hate crimes have been on the rise.⁸⁸ A similar trend has been noticed for anti-Semitism in Europe as well.⁸⁹ Therefore, even if hate speech regulation were to be justified in principle, we should reject its use in practice due to its ineffectiveness, its threat to individual liberty, and its danger to disproportionately harm minority groups.

Section IV: The Democratic Censor

The final hypocritical censor I will consider anchors its justification for censorship in the value of autonomy. In the first part of this chapter I considered and rejected a shallow conception of democracy founded on “majority rules” but this is not the only conception of democracy. In fact, we should pursue a richer account of justified democracy. This kind of democracy is founded in autonomy. The government must reflect the will of the people because individuals have autonomy which grants a sphere of influence that the government and other individuals are not able to touch.

⁸⁶ Strossen, 137.

⁸⁷ Ben Quinn, “Hate Crimes Double in Five Years in England and Wales,” *The Guardian*, October 15, 2019, <https://www.theguardian.com/society/2019/oct/15/hate-crimes-double-england-wales>.

⁸⁸ Rachel Savage, “Rising Populism Stokes Homophobic Hate Speech across Europe,” Reuters, February 4, 2020, <https://www.reuters.com/article/us-europe-lgbt-rights-trfn/rising-populism-stokes-homophobic-hate-speech-across-europe-rights-group-idUSKBN1ZY0X3>.

⁸⁹ Eva Cosse, “The Alarming Rise of Anti-Semitism in Europe,” Human Rights Watch, June 4, 2019, <https://www.hrw.org/news/2019/06/04/alarming-rise-anti-semitism-europe>.

The authority of government actors and fellow citizens in this model are constrained by principle. Even if it is agreed that democracies derive its value from autonomy there is a divergence on what kind of autonomy the government can justifiably concern itself with. I argue here that this difference is generated by a dispute over the importance of using coercion to protect positive and negative autonomy.

The liberal censor that I foreshadowed in chapter 1 is concerned with restricting speech when it violates negative autonomy. In short, censorship can be justified when speech unjustifiably interferes with the autonomy of others, like in the theater case where the intentional speech threatens others with injury and death. Virtually everyone agrees that force is justified in repulsing assaults on negative autonomy such as attempted theft and murder. However, the democratic censor adopts a more capacious concern with autonomy. This censor thinks that speech can also limit the positive autonomy of individuals and that limitations on positive autonomy present a principled reason to restrict speech.

Before we go further it will be helpful to elaborate more on the distinction between positive and negative autonomy.⁹⁰ Isaiah Berlin famously made the distinction between positive and negative liberty.⁹¹ He notes that negative liberty is the absence of interference or coercion from others. Negative liberty is “freedom from” others while positive liberty is “freedom to” do what one wants.⁹² These liberties can come apart. Imagine an abjectly poor person in a thoroughgoing capitalist society. The person may have complete negative liberty with no one interfering or coercing him while his poverty completely restricts the life that he wants to live. Berlin notes that

⁹⁰ For clarity, throughout this paper I will be using the terms liberty and autonomy interchangeably. I recognize that there is a vast literature discussing these two different terms.

⁹¹ Isaiah Berlin, *Four Essays on Liberty* (Oxford University Press, 1969).

⁹² Berlin, 22.

the importance of positive liberty comes from the desire for self-mastery where individuals are “moved by reasons, by conscious purposes, which are my own, not by causes which affect me, as it were from outside.”⁹³ But limiting forces from the outside can go beyond interference from others. A lack of money can foreclose one’s aspirations like coercion can. While every infringement on negative liberty limits one’s positive autonomy to at least a certain extent, it is not true that all deprivations of positive liberty result in losses of negative liberty. If Sue has her house destroyed by a hurricane she loses a significant amount of positive autonomy while still being completely “free from” her fellow citizens.

Having separated the concepts of negative and positive liberty, we must now turn to the debate over the role the government should play in realizing these values. Remember, most people think the government is justified in protecting the negative liberty of individuals. As outlined in chapter 1, the government is justified in responding with force when someone wrongly interferes with the actions of others. However, some go further than this and claim that government also has the responsibility to promote the positive autonomy of individuals. In other words, the role of the government is to enhance the freedom of individuals so that they have the freedom to do things. For instance, some may argue that government cannot neglect those who are trapped in poverty and would otherwise starve in a capitalist system. Providing food, money, or other types of welfare are means to improve the positive autonomy of that person. There are three different standards that can be adopted that the government has the duty to meet for its citizens with respect to positive autonomy.

⁹³ Berlin, 23.

1. No Standard

- a. This position asserts that government has no duty to promote the positive autonomy of its citizens. The government only has the duty to protect the negative liberty of its citizens.

2. Threshold Standard

- a. This argument claims that the government does have a duty to promote the positive autonomy of its citizens. However, this approach denies that government has the responsibility to maximize positive autonomy and instead contends that governments only have the duty to raise their citizens to a certain threshold of positive autonomy.

3. Maximization Standard

- a. Proponents of this claim argue that government has the responsibility to maximize positive autonomy. In short, in every instance where the government is able to promote positive autonomy on balance it has the duty to do so.

Of these three standards, the first and the third will likely strike the reader as very implausible. The first standard might be rejected because in some cases we hold that individuals have duties to support the positive autonomy of people. Imagine that you are walking by a person who is drowning. You have the ability to pause your walk and easily save that person's life.⁹⁴ The drowning person is not suffering from a lack of negative autonomy. To their chagrin, no one is presently interfering with respect to how their life is going. The person faces the threat of the water permanently depriving them of positive autonomy. If the reader thinks you have a duty to intervene in the

⁹⁴ This case is drawn from Peter Singer's influential article. Peter Singer, "Famine, Affluence, and Morality," *Philosophy & Public Affairs* 1, no. 3 (1972): 229–43.

situation and save the person from drowning then you reject the first standard for individuals and think there are at least some duties of aid for the positive autonomy of others. If you are one to think that individuals have no duties of aid to others or that individuals do but governments do not then you will reject the positive autonomy argument from the start. However, I think that a large share of readers will not take this position so I will explore the assumption that individuals and governments have at least some duties to promote the positive autonomy of others.

Having assumed there are at least some positive autonomy duties I will now consider the remaining two approaches. The second standard says that you have a duty to promote the autonomy of others up to a certain threshold. The person drowning in the water may be facing threats to their positive autonomy from drowning and not having enough money to pay for their college education. One may say that you have a duty of aid up to a particular threshold where an individual is required to pull the person out of the water but they are not required to fund their college education. Here the implausibility of the maximization standard rears its head. If an individual has the duty to maximize autonomy wherever possible then they would have the obligation to both save the person from the water and personally fund their college education if that would maximize positive autonomy. Most do not think duties of aid would extend this far and would stop short at some lower threshold. However, if we refer back to Chapter 1, the position that government has the duty to maximize positive autonomy wherever possible can be universally adopted and is not contradictory. Therefore, I will consider this idealized censor in chapter 3 under the name of the Maximal Autonomy Censor.

Having laid the foundation for thinking about the different roles individuals and governments could take in thinking about their duties about positive aid, we can now discuss the implications these standards can have for thinking about speech and censorship. Speech can certainly enhance or diminish the positive autonomy of others. If governments have duties of promoting positive autonomy then this could imply a duty to censor or require certain forms of speech. A basic example concerns compliments and criticisms. Complimenting others can improve their self-esteem and increase their confidence and performance in projects they want to realize. Similarly, harsh criticism can shatter self-esteem and diminish the positive autonomy for individuals to accomplish what they want to do.

If the first standard is right that government has no duty to promote the positive autonomy of its citizens then the positive autonomy argument would be unable to generate a justification grounded in duty to censor speech. If the maximizing autonomy standard is correct then government would have the duty to censor speech whenever speech undermines positive autonomy on balance. Likewise, this argument would also imply that government has the duty to use coercion to mandate speech to maximize the positive autonomy of others.⁹⁵ Between these extremes, many theorists try to thread the needle and argue in favor of government having a duty to promote the autonomy of its citizens to a certain level. The hope is to provide a principled justification of censorship that would allow government to enhance positive autonomy without committing to an implausible moral principle that is illiberal in its application. I will first outline this type of positive autonomy argument and then respond that it is unable to maintain a principled commitment to speech censorship.

⁹⁵ I will expand on more detail in chapter 3 as to why I think this argument would be committed to this position.

I will argue that even if one wants to reject my threshold argument that a threshold argument still cannot provide a stable ground for censorship of speech. The reason is that individuals and governments can help people reach a positive autonomy threshold in a variety of ways that go beyond simply speech. It is not true that censoring speech is the *only* way to promote positive autonomy, even if it is a way. Thus, promoting positive autonomy to a certain threshold would be an imperfect duty that fails to require certain patterns of speech to be consistently performed. Rationally justified censorship according to moral principle, as discussed in Chapter 1, would require a principle that is consistent across all rational agents and one that does not contradict other moral principles. Coercive laws cannot principally require conduct that individuals have imperfect duties to perform. This is because the laws would end up using coercion to require conduct that individuals in some cases have no duty to adhere to. It is categorically wrong to use coercion to impinge on the autonomy of individuals when they do not violate any moral duty. Coercion should be used only as a last resort when all other plausible non-coercive options have been exhausted. Thus, the threshold positive autonomy justification cannot be extended plausibly on principle and should be rejected.

An influential threshold argument for promoting autonomy through censorship is offered by the legal scholar Cass Sunstein. Individuals like Sunstein think the government has a role to play in the promotion of positive autonomy. Ultimately, he thinks that a society where authority is vested in a democracy full of fully autonomous agents is the best one. It is not enough that government simply protects individuals from violations of their autonomy. Instead, government should go further and also make sure that the capacities of individuals are enhanced so they are better able to achieve what they want as fully autonomous agents. Sunstein's support for positive

autonomy focuses on influencing how individuals make choices in order to improve their lives. Actors in this view are justified in engaging in interventions for the sake of the positive autonomy of the individual. Sunstein links this view up to his view of speech in a democratic society. If speech is integral to autonomy then the government should be strongly invested in protecting it. Sunstein explains that “The right to free speech is hardly in tension with democracy, it is a precondition for it.”⁹⁶ A robust conception of democracy is concerned with autonomous individuals exercising their capacities and this cannot be done without speech.

Sunstein also sees drawbacks with complete freedom of speech and sees a destabilizing effect such freedom can have on realizing full autonomy. Rae Langton agrees with this viewpoint and references the harms to autonomy specifically that comes with pornography. Langton argues that certain forms of speech, like pornography, carries with it illocutionary authority that in certain contexts functions as an “illocutionary act of subordination” by promoting a narrative of women possessing inferior social status.⁹⁷ This position focuses on the contention that speech can have power and that power can be directed negatively against the autonomy of others. Many forms of pornography carry messages of social domination that end up silencing or marginalizing the voices of women as agents of full and equal standing to men. Langton notes that “If pornography subordinates women, then it is not in virtue of its content but of its authority that it does so.”⁹⁸

Democracies decide through deliberation. The benefit that democracies have provided humanity is a mechanism for replacing the making of decisions through force

⁹⁶ Cass Sunstein, *Democracy and the Problem of Free Speech* (Free Press, 1995), 121.

⁹⁷ Rae Langton, “Speech Acts and Unspeakable Acts,” *Philosophy & Public Affairs* 22, no. 4 (1993): 302.

⁹⁸ Langton, 313.

with persuasion and deliberation. The practice of deliberation is important for democracies because of its respect for autonomy. The process of true deliberation is never coercive and the very act of deliberation focuses parties towards consensual agreement that will best promote the positive autonomy for all parties involved. Deliberation is where Sunstein focuses on promoting autonomy. He does not support widespread government censorship because of its threats to individual autonomy. However, Sunstein does take issue with complete freedom of speech because of its tendency to degrade democratic discourse.⁹⁹ One only has to look at social media and their Twitter feed to get the sense that the anonymity, vicious back and forth, dunking on others, and character attacks may not be furthering the ideal of democratic discourse. Sunstein thinks that government has the obligation to protect a necessary threshold of democratic discourse so that positive autonomy can be promoted at least up to this threshold. Sunstein's approach is not maximizing and he does not think that government should use coercion to maximize positive autonomy for individuals across the board. His concern is with achieving a basic standard of deliberation and that sometimes complete freedom of expression can pose an obstacle to this goal.

Sunstein proposes a "New Deal for Speech" that seeks to promote the positive autonomy of individuals. In this model, government policy is directed to "reinvigorate processes of democratic deliberation, by ensuring greater attention to public issues and greater diversity of views."¹⁰⁰ Speech that undermines democratic deliberation should be limited to increase the overall level of democratic deliberation to a basic standard. This approach would ensure that all individuals have access to a basic level of positive autonomy. This proposal draws important comparisons to Franklin

⁹⁹ Sunstein, *Democracy and the Problem of Free Speech*.

¹⁰⁰ Sunstein, 119.

Roosevelt's economic New Deal. This view of government saw unfettered capitalism as having failed to protect an adequate amount of positive autonomy to individuals in the economic sphere. Accordingly, government implemented policies, like Social Security, to raise individuals to a certain threshold of economic freedom. The government did not attempt to maximize positive autonomy but instead tried to set and meet a threshold standard it believed all individuals were entitled to. Sunstein's New Deal for Free Speech works in the same manner but replaces its concern for economic freedom with political freedom relating to democratic deliberation. Even if a political order with robust deliberation that promotes positive autonomy is in the interest of everyone that does not mean that complete freedom will lead to this outcome. Isaiah Berlin notes that there can be a disconnect between what individuals "actually seeks and chooses" and in some cases coercing someone for their own good may "be for my benefit; indeed it may enlarge the scope of my liberty."¹⁰¹

Sunstein and Langton believe that democracies should be committed to the principle that citizens should be able to decide and deliberate on the actions of government in full and equal standing. There are two important clauses in this principle and those are the terms full and equal. The second clause says that citizens should be held in equal standing. This is an important principle and all citizens should be formally equal before the law. However, it is more difficult and dangerous to enforce equality of outcomes between citizens when it comes to speech. An equality speech censor can only remain consistent by fully committing to the principle of equality in a universally coherent manner. In the next chapter I will elaborate on the drawbacks of the approach this kind of censor takes. The second important clause is

¹⁰¹ Berlin, *Four Essays on Liberty*, 25.

the term “full” when we refer to democracy’s desire to hold people in full standing as citizens. Sunstein and Langton focus on this word as it relates to the autonomy of individuals. They argue that speech can undermine deliberation and subordinate others through the illocutionary authority of certain kinds of speech. Individuals must have a requisite amount of autonomy to participate in full standing within a democracy. This view proposes that it can sometimes be justified to limit free speech in the name of autonomy. If a robust conception of autonomy is the primary value in a democratic society then it would be permissible to limit speech to reach a basic threshold of positive autonomy for everyone. Speech would be censored in the name of speech. Limitations on freedom would be enforced in the name of freedom.

A major problem with this argument is that promoting positive autonomy cannot lead to a principled and categorical justification for censoring speech unless it is maximizing. I will argue that if one accepts that there is a duty to help others reach a threshold of positive autonomy then it would be an imperfect duty because that threshold can be satisfied in a variety of ways. It is not true that censoring speech is the *only* way to promote positive autonomy, even if it is one way. Coercive laws cannot principally require conduct that individuals have imperfect duties to perform. This is because the laws would end up using coercion to require conduct that individuals in some cases have no duty to adhere to. I assume that it is always wrong to coerce individuals by limiting their speech when they do not violate any duties they owe to others. Coercion should be used as a last resort and should not be exercised if there are non-coercive options that reasonably achieve the same outcome.

If the justification for censoring individuals in the name of reaching a threshold of autonomy cannot be extended on principle without contradiction then it should be

rejected. I will argue that only two routes avoid such a contradiction in principle. The first route is to reject the threshold argument proposed by those like Sunstein. Another approach is to resolve the contradiction in favor of a principle that justifies censorship in the name of maximizing the autonomy of individuals. This approach would resolve the contradiction by arguing that individuals have a duty to maximize the positive autonomy of others. I will review this approach in more detail in the next chapter in my discussion of the Maximal Autonomy Censor.

I have mentioned that there are different ways of interpreting the clause that citizens should be “full” members of a democratic society. One way is the maximal approach that says that a full members of a democratic society is someone who is as autonomous as possible. Another view is that “full” refers to citizens achieving an adequate level of positive autonomy to be a full participant. This latter interpretation is a threshold approach to positive autonomy and is what I will consider now. Let us assume for argument’s sake that this view is correct and that citizens are morally entitled to a certain level of positive autonomy when they are a member of a democracy. Nevertheless, the positive autonomy of citizens can be elevated or undermined in a multitude of fashions. For instance, a person can have their positive autonomy enhanced through job offers, money, words of praise that boost self-esteem and so on. Positive autonomy can correspondingly be reduced through the revocation of job offers, loss of money, or words that harm one’s self-esteem. It is possible to add to the short list that I have offered to produce an almost infinite array of actions that influence positive autonomy. Thus, while speech does have an impact on positive autonomy it is only one factor of many that can support or hinder individuals in meeting their autonomy threshold.

If individuals only have to reach an autonomy threshold and if that threshold can be reached in various ways then that would at most produce an imperfect duty for individuals and government actors. Kant famously made a distinction between perfect and imperfect duties. Perfect duties are those that must be performed on every occasion. Murder is something that you are always expected not to do and it is a perfect duty. In contrast, charity is an imperfect duty because individuals might be said to have the duty to be sufficiently charitable without having the obligation to give to charity every time they have the chance. Sunstein's autonomy threshold argument is better captured by the second class of duties that requires individuals to meet a certain threshold but it does not mandate that individuals do *everything* required to elevate or not diminish the positive autonomy of others. If we accept this though we still should not say that it offers a principled justification to censor the speech of others. Barbara Herman emphasizes that when agents will they are "committing ourselves to will the necessary means, so long as we do not abandon the end."¹⁰² What Herman is saying is that it is only required to will the necessary means to meet an end and if an end does not require performance on every occasion then it would be an imperfect duty. As we have already established, certain kinds of speech are not a necessary condition for realizing a threshold of autonomy for others. Therefore, speech responsibilities should be considered as imperfect duties.

There are many routes to satisfying this autonomy threshold duty if it exists so no agent should be forcibly committed to a particular mean when there are other means available to achieving the same end. As an example, imagine there is an imperfect duty of charity. A rich person walks past several street beggars one day

¹⁰² Herman, *The Practice of Moral Judgment*, 53.

without giving them charity. However, in a year the rich person will donate a million dollars to an organization that feeds the hungry. It seems correct that the rich person plausibly satisfies her charitable obligations even though she does not give to the beggars. Now let's construct the same case but with a threshold duty of autonomy as an imperfect duty instead of an imperfect duty of charity. Let's say that the rich person has an obligation to help individuals reach a certain level of autonomy. She is not a kind person with her words. She says several hurtful things to some people which diminish their positive autonomy, she fails to say things that increase their positive autonomy, and she doesn't really assist others in cultivating their full deliberative capacities as democratic citizens. These slighted people are similar to the street beggars who are supported by the rich person. However, assume that the rich person at the end of the year donates a million dollars to an organization that feeds the poor and dramatically improves the positive autonomy of individuals. It is likely that the rich person has satisfied her overall imperfect duty to support individuals up to a certain level of positive autonomy. If the rich person can satisfy her duties to support the positive autonomy of others then it appears that there is no claim the government would have on using coercion to force the rich person by law not to engage in speech acts that are beneficial or at least not detrimental to the positive autonomy of others.

At this point one may argue that losses to positive autonomy are special. One may think there are not duties to raise the positive autonomy of others but speech that *diminishes* the positive autonomy of autonomy can be restricted. However, this argument would run right into the same implication that comes from adopting a threshold principle rather than a maximizing one. The adopted principle could be that individuals have the obligation to not speak in ways that diminish the positive

autonomy of others below a certain threshold. This threshold argument though generates the same type of imperfect duty discussed because individuals could diminish a person's positive autonomy through speech and then correspondingly raise it via other means to reach the threshold. The way to escape is to either say that individuals do not have a duty to others with respect to their positive autonomy or that one may *never* act in a way that diminishes the positive autonomy of others on balance.

What has been discussed so far relates to an individual's moral duties. However, the government's authority to coerce others does not neatly align with individual moral duties. For instance, adultery is an immoral act and is a breach of one's moral duty to remain faithful to their spouse. Nevertheless, most readers would say that the state lacks legitimate authority to imprison adulterers or use physical force to make them remain faithful to their spouse. In chapter 3 I will argue that the state only has authority to use coercion in order to protect individuals from violations of their negative autonomy rights. Adultery is a harm to the cheated spouse's positive autonomy and severely compromises their well-being. But not all losses to well-being are actionable when it comes to state authority. The core problem with Sunstein's argument is that it contends that government force can be permissibly used against speech that negatively affects the positive autonomy of others. Accepting a principle that coercion can be used to redress harms to positive autonomy would be a stunning and dramatic moral principle if it were to be true. To highlight the difficulties with accepting this principle consider the following case:

The Gay Chef:

Paul is a gay chef in both senses of the word where he is attracted to men and is a very happy chef. He has invested his life in cooking and it embodies one of his deepest passions in life. One bad day Paul encounters two different disconcerting types of speech. The first instance occurs when on his way to work a random stranger yells homophobic slurs at him. He shudders but brushes off the slurs and enters his home. When Paul enters his home his husband is waiting for him. His husband tells him that he hates his cooking and that Paul is a terrible chef who has no promising chance of success in life. Paul's husband further notifies him that he lied about liking Paul's cooking to preserve his feelings but that he would not do so any longer.

In terms of losses concerning positive autonomy, Paul would be much more negatively affected by the speech of his husband than by the knowledge that a random stranger dislikes gay people. For Paul, the homophobic slurs are likely to ruin his morning while the speech of his husband is likely to end his career. If we accept the principle that the state can use coercion to prevent losses of positive autonomy then the state would have a more actionable and keen interest in using coercion against Paul's husband for his "crime" than against the random stranger. If we accept Langton's argument that speech can be regulated due to its capacity in certain forms for "illocutionary subordination" then the speech of Paul's husband would have enormous power to subordinate Paul. The reason is because Paul's husband occupies a position of trust which grants his speech more authority than the speech of others. What we see from this example is that speech certainly has power and individuals hold positions of authority where their speech can harm or fail to elevate the positive autonomy of others. Yet, the state should be silent when it comes to the relationship

between citizens with respect to positive autonomy. The state can only use coercion when the negative autonomy rights of individuals are threatened. Democratic principles obligate the government to respect citizens fully and equally when it comes to their negative autonomy. The domain of positive autonomy is a realm that lies beyond the authority of the state to forcibly intervene. The only way to breach this realm of positive autonomy would be to accept a principle that coercion can be justified against others in the name of positive autonomy. This principle could be consistently maintained but that fully extended principle is not a liberal one.

I have argued that accepting the coercive principle for positive autonomy would be deeply flawed. First, the threshold argument, even if its core assumption about coercion and state authority is true, would not be able to justify principled speech censorship. This is because autonomy can be elevated or diminished in a variety of ways. Speech is not necessarily connected to the overall fate of one's positive autonomy so at most speech would demand an imperfect duty to respect the positive autonomy of others. Coercive laws can overlap with strict duties because strict duties must always be performed. However, imperfect duties do not always require specific actions so censorship could not be used against particular speech acts. The only way to do so would be to accept the moral principle that coercion can be justified against individuals who do not do anything to violate their duties towards others. Adopting a maximal principle, like people have the duty to never do anything that diminishes the positive autonomy of others, would avoid this first issue but dive right into another. I will discuss my objections to this principle more fully in the next chapter but the principle, in brief, is wrong because the state should not have the authority to use coercion to enforce losses or gains in positive autonomy.

Finally, there is a practical objection to this kind of positive autonomy argument offered by Sunstein. He argues that democracies should regulate speech in the overall name of democratic deliberation and discourse. However, deliberation is a fuzzy and somewhat arbitrary standard that is liable to be hijacked by those in power to serve their ends. Would a dissident who loudly and vehemently criticizes the government and the status quo in a public park be jailed because he is undermining the “deliberative status of his fellow citizens?” Would a rival political party have their ability to raise funds curtailed because money would “corrupt the deliberative process of democracy?” These questions are raised to show that one’s view of “proper” democratic deliberation is a very open question and that it is practically dangerous to let government coerce others on the basis of open questions rather than on the basis of closed and more clearly defined moral principles.

Another practical problem is that Sunstein would have government officials make determinations about when the state can use coercion to enforce standards of democratic deliberation. Sunstein points out that there are shortcomings with how people actually engage in democratic deliberation. However, Sunstein’s arguments against the ability for people to effectively engage in discourse are the same kinds of arguments that can be leveled against public officials. As Christopher Freiman notes, public officials are people so if we cannot trust people to make smart deliberative choices then we also should be skeptical about trusting public officials to make smart decisions coercing people to make smart deliberative choices.¹⁰³ A further argument is required to show that public officials possess an elite skill set to overcome the shortcomings that the common folk engage in. Furthermore, even if such elites were

¹⁰³ Christopher Freiman, *Unequivocal Justice* (Routledge, 2019).

shown to exist many governments do not have those elites installed in the right positions of authority to properly make decisions. Therefore, Sunstein's argument is flawed because he has not supplied an argument that justifies elitism on the part of public officials. Until such an argument is adequately given we should not trust public officials to enforce this moral principle just like Sunstein does not trust the people to engage in proper democratic deliberation on their own.

Conclusion

This chapter surveyed four different types of censors which attempted to offer frameworks for censoring speech while still fostering a broadly tolerant liberal democracy where people are able to broadly express themselves. These censors agree that censorship is justified in cases like the Theater Fire Case where it is necessary to preserve the negative autonomy of individuals. These censors also want to carve out further exceptions for censorship that will foster liberal values. Nonetheless, in the words of Joel Feinberg, these censors are avoiding adopting principles that would bring "wholesale and intuitively unwarranted legal interference."¹⁰⁴ As I explained in chapter 1, a hypocrite is a person who acts and believes in two simultaneously contradictory principles. The failure to resolve contradictory principles is a failure of rationality and all censors have duties of rationality to act on the basis of universally coherent moral principles. This chapter has worked to show that the four "liberal" censors considered in this chapter are hypocritical because they commit to a liberal conception of government while also supporting principles of censorship which have extraordinarily illiberal implications. I have argued that all four of these censors offer

¹⁰⁴ Feinberg, *The Moral Limits of the Criminal Law: Offense to Others*, 27.

extensionally inadequate principles that hold different forms of speech to different standards to preserve the illusion of liberality.

All four of the censors in this chapter try to censor speech according to liberal intuitions. The cultural relativism censor wants to avoid imperialistically imposing values on other cultures so they argue that each culture should be able to determine its own speech standards. Accepting the cultural relativist censor would require stripping morality of its authority or promoting a kind of “majority rules” democracy where majorities get to decide their speech standards. Neither of these options are liberal because they would limit the moral resources of minority groups or endanger them to unjust majoritarian censorship. Not many people who say things like “Europe gets to have their speech standards and we get to have ours” actually want to commit to a principle that allows untrammelled majoritarian power. Certain objective moral principles are thought to prescribe limits on the authority of governments.

The Humanitarian Censor attempts to censor speech that is offensive to others. They recognize that speech can provoke very negative responses in people and strive to protect people from uncomfortable states. However, universalizing a principle of offense, even in its modified form, would offer a principle of censorship that would allow the subjective perception of offense to function as a reason for censoring the speech of others. Additionally, there is no non-arbitrary basis for determining which kinds of offense suffered by individuals get to count when it comes to making censorship decisions. A way to make The Humanitarian Censor operate according to an extensionally adequate moral principle would be to become a Utilitarian Censor which makes decisions according to the utility lost or gained across persons from

speech. But, as we shall see, the Utilitarian Censor is not a liberal one and accepting this position would have far reaching implications for state censorship.

The third censor is the Hate Speech Censor which argues that certain forms of speech threaten the dignity of persons by damaging social assurance. The Hate Speech Censor wants to protect people from hate speech. Nonetheless, the definition of hate speech is extensionally inadequate. It does not make definitional sense to restrict limits on hate speech to only minority groups because this is an arbitrary decision and what a minority group is depends on the frame of reference one takes. Additionally, minority groups can fluctuate over time. The Hate Speech Censor could avoid this issue by adopting a moral principle that uses censorship to equalize social standing across persons. This is the Radical Egalitarian Censor which will be considered in the next chapter. However, the Radical Egalitarian Censor would not be a liberal censor and would require government censorship on a large scale to achieve its intended goal. A Hate Speech Censor who wants to protect only certain groups from hate speech is hypocritical because it holds groups to different standards which can be extended to additional groups and keeping the restrictions where they are is an arbitrary moral decision.

The final censor is the Democratic Censor which seeks to improve liberal democracy by censoring speech which erodes democratic discourse or relationships that preserve a healthy democracy. This view argues that we should promote a certain threshold of positive autonomy for individuals. The problem with this view is that positive autonomy can be promoted in a variety of ways beyond speech. Individuals at best would have an imperfect duty to promote positive autonomy and speech would not be a necessary condition to meet this imperfect duty. It is categorically wrong to

coerce people through censorship when it is not necessary to do so and there are other reasonably non-coercive ways for doing so. Coercing individuals to meet their imperfect duties would be illiberal and this is exactly what the Democratic Censor proposes. One could avoid the charge of hypocrisy by arguing that promoting positive autonomy is a perfect duty which would mean that speech that diminishes or avoids facilitating the positive autonomy of others would violate a duty individuals owe to others. This is a Maximal Autonomy Censor but this censor imagines an entirely different and intrusive role for government that a liberal society seeks to avoid. Here is the breakdown of the hypocritical censors on the left and their corresponding ideal and extensionally adequate censors on the right.

The Speech Relativism Censor	Majority Rules Democracy
The Humanitarian Censor	The Utilitarian Censor
The Hate Speech Censor	The Radical Egalitarian Censor
The Democratic Censor	The Maximal Autonomy Censor

While the ideal censors will be discussed in the next chapter, there are several important conclusions that have been reached in this chapter. I have argued that it is hypocritical and irrational to censor speech on the basis of community standards, offense, hate speech, or promoting more effective democratic discourse. None of these justifications can serve as a stable moral foundation for censorship. Additionally, committing to any extensionally adequate moral principle that allows censorship for community standards, offense, hate speech, or democratic discourse would require committing to an ultimately illiberal government regime. As established in Chapter 1,

aspiring censors have the duty to be rational and avoid hypocrisy. The aim of this chapter has been to expose four popular models of “liberal” censorship as being hypocritical and illiberal. These censors can resolve the contradiction in their principles in two ways. First, they can be liberal and not censor speech according to community standards, offense, hate speech, or democratic discourse. Alternatively, these censors could censor these kinds of things and commit to various forms of illiberalism. The next chapter will outline the stakes of this decision.

Chapter 3: Choices Among Consistent Censors

Introduction

In the first chapter I made a distinction between moral theories satisfying formal and substantive rational requirements. I argue that all viable moral theories must be universal in principle and must be coherent with all other moral principles. This is the formal framework of rationality. In effect, this framework demands that moral principles are consistent and applied in a way that is non-arbitrary. Nonetheless, there are also substantive considerations at play. There are many principles that formally succeed while being substantively inadequate. The censors considered in this chapter are all consistent and meet the formal standards of rationality. The focus of this chapter will be to delve into the substantive implications of accepting various moral principles concerning censorship.

In Chapter 2 I applied the formal structure of rationality to several censorship approaches that are hypocritical because they fail to consistently apply moral principles in a way that is not arbitrary. All of the censors in chapter 2 market themselves as promoting a version of liberalism but consistently committing oneself to their vision of censorship would be illiberal. All of the censors in chapter 2 can be consistent by deviating from principles of liberalism. This chapter will outline the costs one will have to make to practically commit themselves to these various deviations to make themselves consistent. Ultimately, I will argue that the only consistent and substantively sound theory of censorship that we should accept is the liberal censor. All the other censors are committed to principle that if universalized in their application would seriously compromise important liberal values and result in a

concerning expansion of state power. Here are the consistent censors that will be considered in this chapter:

- The Liberal Censor
- The Utilitarian Censor
- The Radical Egalitarian Censor
- The Maximal Autonomy Censor
- The Truth Censor

An aim of this chapter will be to show that none of these censors can be jointly accepted. Consistently meeting the foundational commitments of these censors puts them in an irreconcilable tension with the others. However, the liberal censor, while it cannot achieve the complete vision of the other censors, will go the furthest in realizing all the best ends and is substantively the most plausible censor.

Section I: The Liberal Censor

We have finally arrived to the position of the Liberal Censor. In this section I will outline the argument framework the Liberal Censor relies on. I will argue that the Liberal Censor is a morally consistent censor and that substantively it is the *only* acceptable principle of censorship.

The Liberal Censor Principle:

Censorship is justified only when it is necessary to stop or remedy violations of the negative rights of an individual and when there are no other non-coercive methods one can reasonably employ.

This principle does not accept censorship in the name of utility, equality, positive autonomy, or even truth as justifying reasons for censorship. The Liberal

Censor roots its justification for censorship in the Kantian duty to respect the agency of others. This position is universally coherent because it applies in all circumstances where negative rights are violated and prohibits all forms of censorship that do not adhere to this principle. The second clause of the statement is added because censorship is always coercive and interferes with the autonomy of others. If one can stop or remedy rights violations without invasively limiting the autonomy of others then that option should be taken so long as it is not an unreasonable approach.

It will be helpful now to investigate the source of negative rights for the liberal censor because this will clarify the limits of state power that the Liberal Censor endorses. Rational actors are special because we have the ability to set ends for ourselves that we regard as worthy of our attention. The act of setting ends for oneself is an exercise of autonomy or the independent and purposeful self-direction of an individual. The source of our autonomy is our rational agency and as a result the requirements of rationality set boundaries on the exercise of autonomy. The first chapter elaborated why rational agents have a duty to act rationally so I will not repeat those arguments here. Acting in a rational manner demands consistent action. It is inconsistent for a rational agent to use her autonomy to unjustifiably interfere in the autonomy of others because autonomy is the source of the value of the agent's action so infringing on the autonomy of others eviscerates the source of value for the action. One cannot simultaneously claim an entitlement to act in a way that destroys the source of the entitlement itself.

The fundamental right to autonomy, or freedom, is derived from the duty all rational agents have to respect the source of all value: the independence of rational agents to set their ends in a way that is consistent with the independence of all other

rational agents. For Kant, “an action is *right* if it can coexist with everyone’s freedom in accordance with a universal law.”¹⁰⁵ Therefore, an action violates a right if it fails to respect the freedom of others through the violation of a universal law. As Arthur Ripstein explains, Kant’s conception of state authority is constrained by the right individuals have to be independent, or more simply, to be the master of their own affairs.¹⁰⁶ He notes that for Kant the “right to independence entitles each person to use his or her means to set and pursue his or her own purposes consistent with the entitlement of others to do the same.”¹⁰⁷ This right of independence or freedom is the starting point for all other questions or considerations relating to political authority.

If individuals have the right to be autonomously independent from others then it raises a question of when punishment can be justified. Punishment is focused on limiting the actions of others through coercion, and censorship is one form of punishment. The state’s authority to punish is constrained by the rights of individuals. The state has a duty to be rational and rationality demands that punishments be applied in a universally coherent manner. The right individuals have to autonomy limits the application of punishment. I argue that the Liberal Censor can consistently censor speech that interferes with the agency of others with respect to their negative autonomy to be free from unjustifiable interference. Tamar Schapiro points out that violating the autonomy of others “amounts to a refusal to reciprocate within a scheme of shared thought and action.”¹⁰⁸ This scheme of shared thought and action is the practice of making the independence of each rational agent consistent with the independence of all other rational agents. When someone infringes on the

¹⁰⁵ Ripstein, *Force and Freedom: Kant’s Legal and Political Philosophy*, 13.

¹⁰⁶ Ripstein, 29.

¹⁰⁷ Ripstein, 50.

¹⁰⁸ Schapiro, “Kantian Rigorism and Mitigating Circumstances,” 50.

autonomy of others they remove themselves from a shared system of rational cooperation where they are entitled to independence. As Ripstein explains, “whenever someone acts in a way contrary to right, others are entitled to constrain the wrongdoer’s conduct. Such constraint is not an interference with freedom; it is the hindering of a hindrance to freedom.”¹⁰⁹ The justification of punishment rests in the “victim’s antecedent right to be free.”¹¹⁰ Thus, I conclude that punishment, and by implication censorship, is justified to prevent or remedy the unjustified violations of autonomy.

Now we can turn to discuss what kinds of actions constitute unjustified violations of autonomy that can warrant censorship. In the first chapter I developed the Theater Fire Case where a person intentionally and falsely yells fire to start a panic. This is a case where censorship of the individual’s speech would be justified because the speech interferes with the independence of others by making them vulnerable to death or serious injury. The independence of the individual falsely and intentionally yelling fire cannot coexist with the independence of individuals to be free from bodily injury in the theater and cannot be rationally willed into universal law.

I also argue that speech can be censored when it defames the reputation of another. Consider the following case:

The Defamation Case

Bill is angry at his colleague Bill and they do not have a productive working relationship. One day Bill decides to express his anger by writing an email to the entire company that

¹⁰⁹ Ripstein, *Force and Freedom: Kant’s Legal and Political Philosophy*, 27.

¹¹⁰ Ripstein, 55.

accuses Bob of being a rapist. Bill knows that this claim is false and has no evidence to back it up. He sends the email anyways and Bob is subsequently fired.

In addition to the Theater Fire Case, this case also highlights another case where censorship is justified. Bill's speech is false and leads to extreme reputational damage for Bob. The loss of reputation for Bob strikes directly at his autonomy. The reputational loss is an extremely punitive social sanction that is imposed on Bob in a way that interferes with his ability to act independently and receive a fair hearing. I argue that Bob's rights were violated by Bill's defamatory speech. However, defamation can be the target of censorship or punishment only when the speech is recklessly made without regard for the truth or when it is intentionally false. Individuals have a right to their reputation being secured against reckless or intentionally false statements but not against statements that damage one's reputation that do not meet this criteria.

Speech that wrongly interferes with the property rights of another is another domain that can be regulated by censorship. When individuals possess property rights they are entitled to limit speech that violates those property rights. Fraud is an example of this principle at work.

The Perjury Case

Bob is on trial for murder. Bill takes the stand as a witness and says intentionally false things. These falsehoods make it more likely that the jury finds Bob guilty

The perjury case involves speech that can be punished by similar reasoning to the defamation case. The court system is designed to provide due process to Bob so that he is subjected to fair procedures before being punished. Bill's perjury distorts the

application of the procedures by stating knowing falsehoods so that Bob's case is less likely to be decided on the basis of certain factual considerations.

The Fraud Case

Bob is in the market for a new car. He comes up to Bill who is selling his car. Bill tells Bob that his car is in safe working condition and has no problems. These are lies and Bill is aware of the car having several mechanical issues he does not inform Bob about. Bob purchases the car where it later breaks down on the highway.

Bob's rights were also violated in this fraud case because Bill's speech manipulated Bob's property rights to the car through false information. Fraud reflects Bob's manipulated understanding of a contractual transfer so that he can no longer make a consensual decision. Bob's property claim in the car is based on his consent to exchanging money for the car. However, the fraud makes it so that Bob was not able to freely consent to the exchange by the false representation of the property.

The Harassment Case

Bill is angry at Bob so he begins following him around all day screaming obscenities at him. Bob is not able to work as Bill is tapping on the glass outside his door holding up signs saying not very kind things.

Bob's rights were violated in the harassment case because Bill infringes on Bob's right to have a workspace free of disruption or intimidation. Bob is entitled to a reasonable sphere of control and that includes his workplace. Harassment that violates rights interferes with an individual's sphere of control in a severe and pervasive way that undermines an individual's ability to go about their lives in reasonable ways. Bill in this case is not held liable because he said something obscene

or offended Bob. Rather, by following Bob and severely and pervasively intruding on his work and commute he wrongly breaches an autonomous sphere of control that Bob has the right to be left alone in.¹¹¹ While this is a very rough sketch of harassment right claims, it does illustrate that individuals have at least certain rights in some contexts to being free from speech based harassment.

The Nuclear Codes Case

Bill has discovered the codes for launching several nuclear weapons. He decides to share these codes and publish them on the internet.

Bill's speech revealing the nuclear codes can be censored because they violate the rights of individuals not to be exposed to an unreasonable risk of harm. Exposing others to certain degrees of risk violates their rights because individuals as independent actors would have their independence compromised by being subjected to large degrees of risk. Just as the state can regulate behaviors like drunk driving in a crowded neighborhood because of the unreasonable degree of risk to others, the state can also censor speech, like releasing nuclear codes, which violates reasonable risk thresholds.¹¹²

The Bullhorn Case

¹¹¹ In public areas, like a park, a harassment charge is much harder to level because Bob's sphere of control extends in a much more limited capacity in a public park. For example, Bob would not have the right to be shielded from controversial statements in a public park but would have a claim against someone coming up close to him and repeatedly shouting at him.

¹¹² It is beyond the scope of this chapter to go into what the actual risk threshold is. The criminal law has a difficult time determining levels of unreasonable risk and it is not appropriate to wade into this debate here. My main point is that certain behaviors are unreasonably risky and exposing others to unreasonable risk violates individual rights. I assume that the nuclear codes case is a clear example of such an unreasonable risk.

Bill goes outside Bob's house late at night and begins shouting loud things at Bob with a bullhorn and disturbs his sleep.

This case is another violation of Bob's rights because Bob is entitled to a certain degree of control as an autonomous agent which is infringed upon by the bullhorn in a context where Bob has a claim to not be disturbed. Bob's claim against Bill is that he wrongly causes him distress at night. However, it is not sufficient that speech can be prohibited merely because it causes distress. Recall in the second chapter why we rejected the Offense Censor because having an individual's subjective experience determine what speech is allowed is not acceptable. Therefore, I think that we should join Judith Thomson in accepting the Distress Thesis that "We have claims against others that they not cause us 'non-belief-mediated' distress."¹¹³ The Bullhorn Case involves non-belief-mediated distress because the wrong is that Bill is noisy at the wrong time and not because his speech causes distress due to the beliefs they engender. The Bullhorn Case is a type of assault through sound waves and that is what can be regulated. Belief mediated distress is an inadequate standard because all sorts of permissible speech can cause distress. For instance, Bill could tell his work colleague Bob that he thinks he is lazy and incompetent at his job. This speech could cause Bob a lot of distress, perhaps even more distress than in the Bullhorn Case. However, a person's right to be free from interference cannot extend to cover speech that causes belief mediated distress because accepting that principle would not create a universally coherent standard where the freedoms of all rational agents can coexist.

Speech is an important way that individuals can express themselves as individuals. I argue that rational agents have the absolute right to speak so long as

¹¹³ Judith Thomson, *The Realm of Rights* (Harvard University Press, 1990), 250.

their speech does not violate the rights of others. One's right to freedom does not include the right to violate the freedom of others. As Ripstein explains, "The right to say what you think does not preclude liability for fraud, or injuring another person's reputation, or falsely shouting 'fire' in a crowded theater when you know people will be trampled, because each of these deprives others of things to which they already had a right."¹¹⁴ In this section I have engaged in a cursory outline of certain kinds of speech that can be censored or subjected to legal liability by the Liberal Censor.

- Incitement which provokes the immediate and foreseeable violations of an individual's bodily or property rights such as in the Theater Fire Case.
- Defamation which damages a person's entitlement to having their reputation not be destroyed without a fair hearing due to false or reckless speech.
- Fraud which violates the property rights of individuals to consent to property transfers without being subjected to misrepresentations.
- Harassment which violates the rights of individuals to be free from certain kinds of severe and pervasive forms of interference in a sphere of control they are entitled to exercise control over.
- Perjury which violates a person's entitlement to not be subjected to coercive force without a fair hearing and due process.
- Risky Disclosure which violates the rights of individuals not to be exposed to an unreasonable risk of harm such as in the Nuclear Codes Case.
- Assaultive Speech which violates rights by causing non-belief-mediated distress such as in the Bullhorn Case.

¹¹⁴ Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy*, 51.

I do not want to claim that these are the only categories of speech that can be censored by the Liberal Censor. However, I do think that these categories are reasonably comprehensive. Justifications for speech censorship that are intentionally left off of this list are justifications for censorship that are based in appeals to local community standards, offense, hate speech, promotion of positive autonomy, utility, equality, or truth. None of these standards offer a universally coherent justification for censorship that is substantively sound. However, allowing individuals the broad freedom to speak has led to enormous gains in social utility, equality, the reduction of hate speech, and the pursuit of truth. The Liberal Censor offers a plausible moral standard that can promote many things of value while maintaining rational consistency.

The first chapter was an argument that rational agents are obligated to be methodologically rigorous when applying and accepting moral principles. One must only apply and accept moral principles which are universally coherent. A moral principle is not extensionally adequate if it cannot be applied to all categories of cases that fall under the principle. In the second chapter I highlighted four common approaches that were “liberal” censors but were hypocritical because they applied principles that were not methodologically rigorous and were selectively and arbitrarily applied to some speech and not other kinds of speech that would be covered by the same principle. As Kant says, a moral obligation only has moral force when it carries with it “absolute necessity.”¹¹⁵ A valid moral principle binds rational agents in all cases of the principle’s application. The Liberal Censor here is methodologically rigorous and grounds the moral principle of censorship in respect for individual autonomy.

¹¹⁵ Kant, *Groundwork of the Metaphysics of Morals*, 6.

Censorship is justified only when individuals are wrongly interfered with through speech that violates the ability for the equal exercise of liberty under a framework accepted by rational agents operating under universally coherent principles. I understand that accepting the conclusions reached in this chapter can be unsettling. Most people are attracted to promoting utility, equality, positive autonomy, and truth and want to prohibit negative things like obscenity and hate speech. However, adopting a principle of censorship to advance these aims would be even more unsettling if those moral principle are extensionally adequate and not applied in an arbitrary or hypocritical manner. Any methodologically rigorous moral principle has to bite some bullets because some fundamental values cannot be simultaneously achieved. Liberty, utility, equality, and truth cannot all be completely respected at once.

Before I turn to the remaining sections which outline the problems with accepting other consistent principles of censorship, I will close with some final thoughts on why we should accept the Liberal Censor. There is something beautiful in the realization of autonomy: the ability of independently directing your ends under your own authority. Rational agents are not suited to be pawns on a chess board but rather are naturally suited to be chess players. Rational agents occupy a position of independent authority where they, and they alone, are in a position to determine what ends they will pursue so long as it is consistent with the rational ends of others. For this reason, Kant is right that we have the obligation to treat individuals as ends in themselves and never as a means to our own ends. Treating others as a means is a fundamental affront to one's most basic status as a rational agent. Treating others as a means assumes unjustified authority over them.

Speech is perhaps the most central way that an individual is able to set ends they have determined for themselves. Humans do not exist as a vacuum within our own minds. We are driven by a primal urge to *communicate* and make the outer expression of our identity consistent with our internal identity. Speech is the most effective medium for facilitating this transfer of identity and a person who is limited in their speech will experience a tension between their outer and internal identities. When a rational agent violates the rights of others they willingly open themselves to interference to remedy the rights violation and restore the balance of a state of equal rational freedom. However, any act of censorship that does not remedy or prevent rights violations uses the rational agent as a tool for some other end that they have not willingly adopted as their own. Thus, any form of censorship that extends beyond upholding the rights of others is a deep and fundamental wrong. If we are to value autonomy at all we must accept the liberal model of censorship and reject all other forms of censorship. I will now turn to several moral theories that reject the fundamental value of autonomy as self-directed agency.

Section II: The Utilitarian Censor

In chapter 2 we reviewed the Humanitarian Censor who censors speech in order to prevent harm being caused to others. However, the difficulty with this approach was that it did not have a metric for evaluating when harms could be permissibly censored. Things like offensive speech can bring both benefits and harms. The Utilitarian Censor offers a universally coherent approach for censoring speech by holding utility as its value and justifying censorship, or refraining from censorship, in the name of the promotion of utility. This censor proposes the following moral censorship principle:

The Utilitarian Censor Principle:

Censorship is always morally justifiable when the act promotes more utility on balance and is never justifiable when it leads to less utility on balance.

While the literature on utilitarianism is vast, I am going to follow the majority of scholars and define utility as subjective well-being or happiness, or the pleasure felt by the creature directly experiencing it. Utilitarians hold the promotion of utility as the paramount moral value and assert that actions are morally right “in proportion to as they tend to promote happiness, wrong as they tend to produce the reverse of happiness.”¹¹⁶ This principle directly connects to an evaluation of the morality of censorship: it is right when it tends to promote happiness and wrong when it produces unhappiness. Rational agents value happiness and that is why utilitarians are inclined to view happiness as *the* guiding moral principle. This censor is methodologically sound because it advances a morally consistent position. The promotion of utility is a moral principle that extends across all rational agents. Furthermore, the promotion of utility is the highest value so there is no contradiction as all other values are subsidiary to this guiding moral principle.

What does the utilitarian position imply for government censorship? This is unclear because the conclusions the utilitarian reaches depends on the facts on the ground based on the actual balance of utility that various forms of speech affect. John Stuart Mill offered the most famous defense of individual freedom of speech against government censorship on utilitarian grounds. Mill argued that censorship could lead to the suppression of truth, the loss of the ability to sharpen and critically analyze

¹¹⁶ John Stuart Mill, *Utilitarianism and On Liberty*, ed. Mary Warnock, 2nd ed. (John Wiley & Sons, 2003), 186.

one's own positions, and the loss of individuality within a more conformist society.¹¹⁷ He thought that utility would be favorably promoted in a society that had very minimal censorship. Mill acknowledged that there were human limitations in determining which approaches would promote utility on balance with speech issues but while we cannot have absolute certainty we can have "assurance sufficient for the purposes of human life."¹¹⁸ The problem with this argument is that it does not provide very secure foundations for rejecting widespread censorship on behalf of the state. Mill lived in a time where the government lacked the technology it does today to more selectively target speech that it wants to censor. Mill's time also lacked the internet and social media where misinformation, insults, and cyberbullying can be instantaneously transmitted across audiences of millions. Perhaps Mill today would no longer think that he had sufficient assurance to reject censorship according to the Utilitarian Censor Principle. The ability for speech to promote or detract from social utility is constantly shifting according to the uses of speech and the evolving technology used to promote direct the speech. For these reasons we cannot assume that the utilitarian position will not reject widespread censorship. It is time to take a closer look at the merits of the principle itself.

One major problem with the philosophy of utilitarianism is with its attempt to make utility a unifying metric across persons. It is very difficult to see how different categories of pleasure and pain can be measured against each other according to a common currency. One person might stub their toe and another might be called a "vile crook" and both of these circumstances cause pain and diminish the welfare of the person experiencing it. However, it is absurd to assign common values to these

¹¹⁷ Mill, *On Liberty*.

¹¹⁸ Mill, 89.

experiences such that a stubbed toe causes -3 utility whereas being called a “vile crook” by a random stranger is worth -6 utility and -13 utility if it is said by your friend. The human mind simply does not make such calculations in this manner. If individuals do not think this way then how can we expect governments, populated by individuals, to make such decisions? Moreover, happiness is a subjective experience where one person’s happiness or suffering cannot be directly experienced by others. The lack of direct experience makes it that much more difficult to make utility a unifying currency that can establish clear standards for making a decision. The best public officials can do is guess with “sufficient assurance” but such guesses run the risk of granting public officials the ability to make more arbitrary decisions in employing censorship across persons that strips individuals of the moral due process they should be entitled to.

Another major problem of the Utilitarian Censor Principle is that even if utility can be appropriately discerned and aggregated across persons it leaves government deferring to individuals whose experiences we do not think should have weight in moral decision making. Remember that the utilitarian defines utility as subjective well-being or happiness. Utility is utility and there are no distinctions made between good or bad utility on principle. Bernard Williams aptly critiques utilitarianism because it could lead to a racist majority dictating racist policies if they are sufficiently numerous and feel subjectively strong enough about discriminating against others on the basis of race.¹¹⁹ This argument can be directly applied to speech issues.

The Racist Society

¹¹⁹ J.C.C. Smart and Bernard Williams, *Utilitarianism For & Against* (Cambridge University Press, 1973).

There is an isolated society which is incredibly racist and espouses white supremacist ideals. The vast majority of the society gains a lot of utility in censoring, suppressing, and subordinating minority groups. Additionally, these people lose a tremendous amount of utility on balance when there are suggestions made that the racists are wrong or bigoted in their beliefs. A lone person of color lives in the society and one day tells the racists that they are wrong and bigoted and that all races should be treated equally. The lone person of color gains a very small amount of utility through their speech and this gain in utility is vastly outweighed by the loss of utility experienced by the racists.

In this imagined society, utility is best promoted by censoring the speech of the lone person of color and deferring to the large utility gains of the white supremacists. Thus, the Utilitarian Censor Principle asserts that there is moral justification for censoring the speech of the person of color. This is blatantly wrong and the liberty for an individual to speak should not cave to the desires of the racist majority, even if they gain more utility on balance. The right to speak cannot be dependent on the raw balance of utility because that would offer justifications for societies like the racist society to censor speech.

Another problem with the Utilitarian Censor Principle is not just aggregated groups of individuals can limit the autonomy of others but also that individuals can substantially limit the autonomy of large groups of people. Robert Nozick challenges utilitarianism by imagining a “utility monster” which gains such a vast amount of utility by consuming individuals that the moral principle of utilitarianism would justify feeding individuals to the utility monster to create enormous utility gains.¹²⁰ There can

¹²⁰ Robert Nozick, *Anarchy, State, and Utopia* (Basic Books, 2013), 41.

also be utility monsters when it comes to speech issues. In chapter 2 I outlined the case of Sue who is the eggshell IRS agent whose self-esteem absolutely shatters when she is called a thief by Frank for collecting taxes. Sue is a negative utility monster because she loses such an enormous amount of utility from being called a thief that it vastly outweighs any utility Frank gains by being able to express his political beliefs. The speech utility monster concept can also be applied to justify compelled speech.

King George

King George absolutely loves feeling superior to everyone else and when people speak to praise him and grovel. In fact, when individuals come before him and say things like “no one can match your wisdom,” even if they don’t believe it, King George’s utility skyrockets. King George gains so much utility from this groveling that it outweighs the losses of utility from individuals who are forcibly compelled to come before him to grovel and say false things.

King George is also a utility monster of a different sort from Sue. He gains so much utility from compelled speech that it overrides the autonomy of individuals when the Utilitarian Censorship Principle is applied. Any purported moral principle that supports such conclusions in principle is a substantively flawed moral principle and should be rejected. Just because a person stands to gain or lose large amounts of utility from speech this does not provide a moral justification in itself for censoring or compelling the speech of others. The ability for people to speak should be decided on grounds apart from utility.

The Utilitarian Censor Principle should be rejected by any person who accepts any of the other censorship principles on a consistent basis. Ultimately, the Utilitarian Censorship Principle is hostile to the principled pursuit of autonomy, equality, and

truth. The Liberal Censor makes censorship decisions based on individual autonomy while the Utilitarian Censor makes these decisions based on the balance of social utility. Granting individuals liberal rights to speak will sometimes lead to losses of utility on balance. Anyone that thinks that utility considerations should not dictate the freedom of individuals should reject the Utilitarian Censor. The promotion of utility and equality are also foundationally at odds. There are cases where maximizing utility comes at the expense of utility. For example, one action could lead Person A and B to have utility gains of 20 and 20 respectively. This would uphold equality. However, another action could lead to Person A and B having gains of utility of 250 and 20. This second action provides much more utility on balance but creates deep inequality.¹²¹ As John Rawls spells out “the principle of utility is incompatible with the conception of social cooperation among equals for mutual advantage.”¹²² Finally, the principle of utility can also come into tension with the pursuit of truth. It has long been said that “ignorance is bliss” and knowledge of what is true can at times make people deeply unhappy and degrade their utility. The truly consistent utilitarian has to be willing when push comes to shove to suppress the truth and lie when an individual speaks the truth and this results in losses of utility on balance. Moreover, Susanna Rinard argues that there are sometimes cases where people can believe things without being directed on the basis of evidence.¹²³ Utilitarians would recommend in such circumstances people should adapt their beliefs, even against the evidence, in order to maximize utility. On a basic level of foundational moral commitments, the utilitarian

¹²¹ The utilitarian would even support actions that would make person B worse off for overall utility. For instance, if another option created changes in utility of 280 and -20 that would be the most preferable option from the utilitarian perspective even though Person B is actively made to suffer for the utility gains of Person A.

¹²² John Rawls, *Justice As Fairness* (Harvard University Press, 2001), 13.

¹²³ Susanna Rinard, “Equal Treatment for Belief,” *Philosophical Studies* 176, no. 7 (July 1, 2019): 1923–50, <https://doi.org/10.1007/s11098-018-1104-9>.

ensor cannot be reconciled with the principled commitments of the liberal, equality, and truth censors. If one chooses any of the other consistent censors then the Utilitarian Censorship Principle must be rejected.

A final reason against the Utilitarian Censor that I will mention is that the pure pursuit of utility is not a desirable end in itself. Utilitarians assert that utility is the core value humans should stake their moral interests in but the challenge of accepting this conclusion is reflected in an ingenious thought experiment offered by Derek Parfit. The Repugnant Conclusion is a thought experiment developed by Parfit which imagines one productive and healthy world on average and a world in misery but with an extremely large population. Parfit notes that “For any perfectly equal Population with very high positive welfare, there is a population with very low positive welfare which is better, other things being equal.”¹²⁴ To illustrate this issue raised by Parfit in more detail let’s turn to the

Meteor Strike

There are two worlds. World A has a small population of only several thousand. These people are all very happy. They all are able to speak freely and they are autonomous agents who pursue their ends with equal respect for the ends of others. World B is a state of appalling misery. It is populated by several billion people who have lives barely worth living. These people are regularly unhappy and sick, they are censored, they are conformists, and utterly lack individuality. However, the small utility of each life barely worth living aggregated across billions means that World B has more total utility on balance than World A. A meteor is hurtling towards these worlds. You have the ability to

¹²⁴ Derek Parfit, *Reasons and Persons* (Oxford University Press, 1986), 388.

direct the meteor so that only one of these worlds is destroyed but you can do nothing else.

So which world do you save? The choice of which world to save reflects a value judgement. It embodies the world you think has more value, because rational agents would not leave a world they think has more value to be destroyed. If utility is truly the governing moral value then the utilitarian would save World B every time. However, if you think that World A should be saved then you are rejecting the foundational utilitarian assertion that utility is the foundational and ultimate source of value for morality. Robert Nozick aptly highlights the issue with considering utility as an end in itself. He imagines an experience machine where we could plug in and experience anything that we want.¹²⁵ The experience machine is a means for achieving whatever experiences that we want. However, if you would not want to live your life through the experience machine then there has to be something other than utility which is of ultimate value. There is something of value to living as an autonomous agent in an imperfect and messy world that leads me to reject the experience machine along with many others. One may be inclined to reject plugging into the experience machine because if one does so then they would no longer be living in an important sense. The Buddhists are right that to live is to suffer so to plug into a machine that obliterates suffering would be to rob oneself of the chance to live. If one does not want to destroy World A or if they do not want to plug into the experience machine then they are declaring that there is something more to life and morality than the pursuit of utility. If one accepts this then they should reject the Utilitarian Censor Principle on substantive grounds.

¹²⁵ Nozick, *Anarchy, State, and Utopia*, 42.

Section III: The Radical Egalitarian Censor

The third consistent censor is the Radical Egalitarian Censor. This censor believes that censorship should be used when speech undermines the equal respect that ought to be accorded to people. Certain kinds of speech can create group based harms that are an affront to the identity of historically marginalized groups. Words have power and can have disparate impacts on different groups of people. The Radical Egalitarian Censor is committed to promoting equality between individuals and groups and is willing to censor speech in order to maintain a state of equality of standing amongst persons. The view is expressed along the lines of the following principle:

The Egalitarian Censor Principle (ECP)

It is justified to censor speech or conduct that undermines an individual or group's entitlement to have their dignity be respected and to have their dignity be accorded equal standing with others.

I will argue that the ECP is a valid moral principle when it is applied procedurally to respect the rational agency of individuals. I will claim that respecting the rights of persons consists in respecting the negative autonomy of individuals so long as they do the same. Censorship of speech that does not infringe on the negative autonomy of individuals is wrong because it fails to recognize the equal claim of others to negative autonomy. This is a type of relational egalitarianism based in procedure: so long as you adhere to procedural constraints on speech based on the negative autonomy of others you get to say what you want. Thus, the liberal censor is a kind of egalitarian

ensor when it comes to procedure but not according to outcomes. There are others who want to stretch the ideal of relational egalitarianism further to also incorporate and enforce the ideal of substantive equality between persons.

Relational egalitarianism is the ideal, as articulated by Kasper Lippert-Rasmussen, that individuals are part of a community where people treat and regard others as equal.¹²⁶ There are different ways in which someone could promote relational egalitarianism as it relates to the exercise of autonomy. The first kind of egalitarianism concerns itself with procedural equality and argues that every person should have the entitlements to exercise their autonomy on equal standing. The second egalitarianism emphasizes substantive equality and states that there should be equality of outcomes between persons in terms of their autonomy. These two approaches meet the ideal of relational egalitarianism but do so in very different ways that can produce very different results. To illustrate this consider two ways to run a blackjack game.

Procedural Equality Approach:

Every player receives chips and they are granted the autonomy to make whatever blackjack bets they want that are consistent with the rules of blackjack. Every player is subject to the same blackjack rules and cheating these rules is not tolerated. Some players make very risky bets and others make conservative bets. Over time, there are large inequalities of chips that emerge between players.

Substantive Equality Approach:

Every player receives the same number of chips. Players have their autonomy of playing blackjack circumscribed so that they cannot engage in bets that would result in chip

¹²⁶ Kasper Lippert-Rasmussen, *Relational Egalitarianism: Living as Equals* (Cambridge, United Kingdom: Cambridge University Press, 2018), 16.

inequalities between players. Alternatively, players can engage in different kinds of bets but at the end of the game chips are redistributed so that all players have an equal chip amount.

What these examples show is that there are radically different ways of realizing the ideal of relational egalitarianism. The procedural equality approach treats and regards everyone as equal when it comes to the exercise of autonomy, so long as that exercise is consistent with behavior that respects the autonomy of others. Inequalities are permitted between people and do not violate the ideal of relational egalitarianism if everyone is subject equally to the same procedures. The second approach is very different. It focuses on equality of outcomes being central to treating and regarding others as equal. It can realize equality of outcomes by either limiting the procedures to produce equal results or by taking action after the fact to nullify differences between people.

The upshot of this discussion is that the Egalitarian Censorship Principle will have very different implications if one adopts the procedural equality approach compared to the substantive equality approach. There is not a clear way to pick between the two approaches because each could be adopted consistently. In this section I will be considering and rejecting the second interpretation of the ECP. G.A. Cohen argues in favor of relational egalitarianism based in outcomes. Cohen is a socialist and thinks that justice requires that members of a community uphold “some principle of equality in the distribution of benefits enjoyed and burdens borne by its members.”¹²⁷ His concern is that luck will result in inequalities between persons and that is unjust. He proposes to correct such injustices in order to make sure that the

¹²⁷ G. A. Cohen, “Robert Nozick and Wilt Chamberlain: How Patterns Preserve Liberty,” *Erkenntnis* (1975-) 11, no. 1 (1977): 6.

effective freedom of others is not dictated by luck or the decisions of others.¹²⁸ Cohen notes that “a socialist need not apologise for being willing to restrict freedom in order to expand it” in the name of equality.¹²⁹ Kai Nielsen also supports this position by claiming that the government should prioritize promoting equality in terms of benefits and harms between citizens before the liberty of individuals.¹³⁰

Cohen’s and Nielsen’s arguments can be used to justify censorship of speech. There are forms of speech, like hate speech, which target groups and reduce the effectual freedom of members of groups. A lot of speech discussed in the hate speech section of the last chapter fall under this designation. Speech like “no loathsome (insert group identity here) are welcome” serve to limit the effectual freedom of members of certain groups by assaulting their self-esteem and diminishing their sense of belonging. Cohen’s argument leads to the conclusion that speech that contributes to inequality between the effectual freedom of individuals can be regulated in the name of equality. A charge that cannot be leveled against Cohen and Nielsen’s arguments is the claim of inconsistency. They start with the premise that the state is justified in using force to remedy inequalities between groups in terms of effectual freedom. Speech can cause disparities in individuals in terms of effectual freedom so censorship can be justified in these cases.

While Cohen and Nielsen’s arguments cannot be rejected on the basis of inconsistency, it should be rejected because it suffers from several substantive defects. The first is that censorship is wrong because it is a form of leveling down to reach equal outcomes rather than by leveling up. The wrong of leveling down instead of up to

¹²⁸ G.A. Cohen, *On the Currency of Egalitarian Justice and Other Essays in Political Philosophy* (Princeton University Press, 2011), 5.

¹²⁹ Cohen, “Robert Nozick and Wilt Chamberlain,” 18.

¹³⁰ Kai Nielsen, *Equality and Liberty: A Defense of Radical Egalitarianism* (Rowman & Littlefield Publishers, 1985).

reach equality is beautifully captured in the short story of Harrison Bergeron by Kurt Vonnegut. In the story, Harrison Bergeron is a talented young boy who is limited by the state in the name of enforcing equality. He is forced to wear impediments that obscure his physical beauty, weights restrict his physical prowess, and sensors damage the expression of his intellectual gifts.¹³¹ The society is equal by having Harrison Bergeron leveled down to the position of everyone else by coercion. Vonnegut, who was a socialist, was making the point that equality can be pursued in a wrong way if it is done through leveling down others. Leveling up others to a position of equality is always preferable than leveling down. Censorship is coercive so it is always a form of leveling down when it is used to promote equality. An egalitarian state always should pursue a leveling up strategy to reach equality than a leveling down. Individuals and groups may have their equal standing harmed by speech but that equal standing can be restored in other ways, such as the government elevating and encouraging the voices of historically marginalized groups. If censorship is not necessarily required to achieve equality, which it is not, then Cohen's egalitarian argument would not support censoring speech if one accepts the premises that leveling up to equality is preferable than leveling down and that leveling up is possible. Thus, censorship is not a necessary condition for achieving the Radical Egalitarian vision so it should be rejected in favor of other leveling up strategies to equality.

Another problem with the Radical Egalitarian Censor is that egalitarianism is fundamentally at odds with the state's ability to censor speech. The radical egalitarian argues that all individuals should be treated equally with respect to their effectual freedom but then declares that state officials are granted special authority that other

¹³¹ Kurt Vonnegut, "Harrison Bergeron," in *Philosophy, Politics, and Economics: An Anthology*, ed. Jonathan Anomaly et al. (Oxford University Press, 2016).

citizens are not to coerce people through censorship. The state is, by nature, a hierarchical organization that claims a special purview to enforce the laws that ordinary citizens cannot. As Tibor Machan notes, “to bring about the requisite equality there must be a group of persons quite unequal to everyone else in their power to impose their will.”¹³² Therefore, the state should be considered an organization that grants its officials extra effectual freedom to enforce equality. Consequently, a truly consistent radical egalitarian cannot promote state authority for censorship because that would fail to treat individuals equally and would in fact be promoting inequality through the maintenance of a system of hierarchy. If the radical egalitarian tries to grant state officials an exception when it comes to equality of effectual freedom then the argument would be committed to a contradiction that would relegate itself to the status of other hypocritical censors in chapter 2. Thus, a consistent radical egalitarian censor would be opposed to state censorship authority on principle. Without state authority for censorship where the state holds a monopoly on censorship then an anarchic state of affairs is produced where everybody can censor everybody in the name of equality. Most people who want to censor in the name of equality would not want to commit themselves to this vision of society. I think most readers will join me in rejecting this attempted justification for censorship as substantively unsound.

Accepting the egalitarian censor would also require rejecting the visions of all the other consistent censors. First, censorship in the name of equality violates the liberty of individuals by using coercion against them when they have not violated the negative autonomy of others. Second, using coercion to promote equality can undermine the utilitarian’s desire to maximize utility. For instance, imagine there are

¹³² Tibor Machan, *Liberty and Equality* (Hoover Press, 2013), xv.

two individuals in a society where utility can be quantified. Prior to censorship, person A has a 100 utiles and person B has 20 utiles.¹³³ This state of affairs is unequal and let us say that person A's ability to speak freely leads to this disparity in outcomes. The Radical Egalitarian censors person A so that both people have 20 utiles and they are substantively equal. However, this destroys utility. The world before censorship had 120 utiles of censorship and forcibly promoting utility by leveling down led to a world with a total utility of 40 utiles.¹³⁴ This same argument also shows that the radical egalitarian can be fundamentally at odds with maximizing autonomy. If you take the example above but replace utiles with some measure of autonomy we see that promoting equality through censorship can result in a world where positive autonomy is destroyed in the name of equality. Thus, the maximal autonomy censor should reject the radical egalitarian censor's vision. Finally, the radical promotion of equality can also come at the expense of truth seeking. Individuals can generate inequality by speaking the truth because the truth can sometimes hurt or diminish the welfare of individuals. If an official can use coercion to limit the liberty of others to speak in the name of equality then it also would suggest that suppressing the truth in the name of promoting equality would also be permissible for the Radical Egalitarian Censor. Therefore, all the other consistent censors, in virtue of their consistency, reject the moral principle underlying the claim of the Radical Egalitarian Censor.

A final argument against this kind of radical egalitarianism is that substantive equality between persons is not a good unconditioned final end. Remember that the

¹³³ I am using utiles to refer to a universal measure of utility for ease of making this argument. I do not think that a "utile" actually exists and that in practice it cannot be calculated in the manner I refer to in this example. A rejection of "utiles" is discussed earlier with the utilitarian censor.

¹³⁴ Note that leveling up person B to the level of person A would create a world with 200 utiles. The pursuit of equality through leveling up rather than leveling down always results in gains to utility.

Radical Egalitarian Censor is willing to censor speech, limit liberty, destroy utility, diminish overall positive autonomy, and suppress the truth in the name of equality so it better be a wondrous final end. But that is not the case. If people are equal in terms of outcomes then so what? If the reply is that equality makes people feel better or reduces resentment then these are arguments for human happiness, not equality, as being the final unconditioned end. I have argued that *procedural* equality between persons is valuable because it respects the independent autonomy of rational actors. But in this argument respecting procedural equality is a means to ultimately respecting the autonomy of rational agents by respecting the equal moral standing of rational agents. Equality can be an important means but I think it has zero independent value as a final end. Some may disagree and I will be unlikely to convince them otherwise because equality for them is final and they are willing to sacrifice freedom, welfare, and truth in its name. However, my arguments in this section are arguments for why individuals who value other ends should reject the central aim of the Radical Egalitarian Censor to regard outcome equality as a final end. One can consistently commit to the principle of equality but that commitment would require sacrificing things of moral value and produce a more authoritarian world that is chillingly reminiscent of the social vision painted in Harrison Bergeron. Some may want to bite this bullet for equality but that is a decision I will not make and I hope most readers will not make either.

Section IV: The Maximal Autonomy Censor

The Maximal Autonomy Censor is a censor who cares deeply about the value of autonomy but thinks that the state should interact with the autonomy of individuals

in a different way than the Liberal Censor. The Liberal Censor argues that the value of autonomy derives sets of rights for individuals which lead to fundamental constraints on action. The Liberal Censor views the state's role as being to protect the rights of individuals from unjustified interference. The Maximal Autonomy Censor agrees with the Liberal Censor in all of these respects but then goes further and claims that the state has a duty to promote the positive autonomy of individuals. Recall that in chapter 2 we rejected the threshold view of autonomy because it could not establish a threshold of autonomy that should be positively promoted in a non-arbitrary way. The Maximal Autonomy Censor avoids this issue by consistently promoting positive autonomy. Similar to how the Utilitarian Censor strives to maximize utility in the world, the Maximal Autonomy Censor works to maximize positive autonomy. When speech serves to diminish positive autonomy on balance this offers clear grounds for censorship if one is searching to maximize positive autonomy. A way to structure the autonomy censorship principle is as follows:

The Autonomy Censorship Principle:

Censorship of speech is justified only when the act of censorship serves to promote more autonomous action on balance.

A major proponent of this approach to autonomy is articulated by Sarah Conly in her book *Against Autonomy*. In this book Conly argues that paternalistic action against individuals is justified when it is necessary to help individuals realize their actual values.¹³⁵ For instance, individuals value their health and want to live healthy lives. Smoking destroys the health of an individual so Conly thinks that the state is justified in paternalistically interfering with the ability of individuals to smoke in order

¹³⁵ Sarah Conly, *Against Autonomy: Justifying Coercive Paternalism* (Cambridge University Press, 2013).

to help them realize the value which they care more about, their health. Conly contends that “Legislation should intervene when people are likely to make decisions that seriously and irrevocably interfere with their ability to reach their goals, and where legislation can reliably prevent them from making those bad decisions, and where legislation is the least costly thing that can reliably prevent them from making these bad decisions.”¹³⁶ This principle can be directly applied to actions involving speech. People have fundamental goals and values but sometimes people are unable to hold their tongue when it matters and this can destroy the ability to reach their goals. Just as smokers reach for another cigarette even when it contravenes the values they set for themselves, people are also drawn to speak. Consider:

The Vindictive Email

Michael highly values personal relationships like most people. He deeply values the intimate relationships that he has with his family, friends, and work colleagues. One day Michael has a very bad day. His family was incredibly annoying, his best friend was mean to him, and his boss was impatient. Michael was annoyed and decided to write an angry email directed at his family, friends, and his boss. The email is very hostile, vindictive, and venomous. It will certainly serve to drive away these relationships that Michael values permanently. Deep down, Michael knows the consequences of his email but chooses to send it anyways. A government censor intercepts the email before it is sent. The government censor reads the email and infers that the email will compromise the ability for Michael to reach the goals that he autonomously set for himself to have close personal relationships. The email is censored and never received by Michael’s family and friends.

¹³⁶ Conly, 72.

If people autonomously choose to pursue healthy lives with close personal relationships then actions like smoking and vindictive emails can compromise these long term values. The Maximal Autonomy Censor supports censorship in cases like the case above where censorship is necessary to prevent losses to an individual's autonomously directed ends on balance. Censoring the email is an act of paternalism, but theorists like Conly think it is not a bad label at all to be paternalistic: our valuing of autonomy writ large should commit ourselves to stopping individuals from undermining their own autonomy. The Autonomy Censorship Principle is rationally consistent because it says that all cases where an individual's action results in the loss of autonomy or the autonomy of others on balance should be restricted by the state.

While the Autonomy Censorship Principle is consistent, it should still be rejected on substantive grounds. If autonomy is to mean anything valuable at all then it must allow the freedom to make mistakes. Smoking and sending vindictive emails may compromise our long term values, but our decision to smoke a cigarette or to send angry emails are still in themselves autonomous choices and should not be restricted by force. If we were to use coercion to prevent people from saying things that endanger their most valuable plans and projects in life then we would be limiting individuals to just speak on the little things in life. As Robert Jackson once said, "Freedom to differ is not limited to things that do not matter much. That would be a mere shadow of freedom. The test of its substance is the right to differ as to things that touch the heart of the existing order."¹³⁷ Decisions to smoke or send nasty emails are actions which touch on the heart of the existing order of the projects of our lives.

¹³⁷ "West Virginia State Bd. of Educ. v. Barnette, 319 U.S. 624 (1943)," Justia Law, accessed April 11, 2021, <https://supreme.justia.com/cases/federal/us/319/624/>.

The realization of these life projects are the culmination of autonomous choices that are made along the way. If guardrails are put up so that we are not able to derail our life projects then the value of these projects are diminished because we have given them less autonomous self-direction along the way.

If one is inclined to accept the Autonomy Censorship Principle then that person will be committing to a far reaching moral principle that extends beyond issues of speech. The principle is that coercion is justified to forcibly prevent individuals from compromising their long term values. This would justify paternalism in a host of cases in our lives. Smoking, drinking too much soda, eating fatty foods, watching too much TV, and speaking ill of our family and friends are all actions that threaten the ability for people to realize their long term projects in life. Someone who accepts the Autonomy Censorship Principle are committed to leaning into paternalism and using force to limit smoking, soda and fat intake, excessive TV watching, and angry comments to family members where it is feasible to do so. In fact, Sarah Conly urges readers to adopt this paternalistic approach more comprehensively.¹³⁸ However, if you are unwilling to accept such wide reaching state paternalism then you should reject the censorship principle for speech that this argument is based on.

There is also an extreme practical difficulty in actually realizing the Autonomy Censorship Principle. The problem is that individuals are usually the best judges of what actions will promote their long term values and projects. The state is unable, yet, to peer into the minds of individuals to see how individuals conceive of the values they have autonomously given for themselves. Acting paternalistically always limits the exercise of autonomy to some extent. If the aim is to promote the exercise of more

¹³⁸ Conly, *Against Autonomy: Justifying Coercive Paternalism*.

autonomy on balance then paternalistic actions are always morally risk because they threaten the loss of autonomy, and perhaps the loss of more autonomy on balance. An added difficulty is that it is unclear whether or not actions done in the present signal long term shifts in values. An autonomous agent, to be truly autonomous, must be able to change her values over time. For instance, someone may value health at one point but over time come to value the pleasure and sociality that comes with smoking cigarettes more than the damage it causes to one's health. The act of smoking may threaten the agent's long term value of health, or reflect an autonomous choice that is consistent with a shift in long term values. There is no clear remedy for how a paternalistic can plausibly determine which choices damage long term values and which signal shifts in values over time. The paternalistic is committed to a kind of elitism that says that the censor knows how to direct the life of another toward their values better than they do themselves. Such elitism must be justified and shown that the accepted elitism can be maintained over time in government. There is the practical concern that government paternalism will lead to more losses of autonomy to individuals on balance. Such practical concerns are reasons to reject the Autonomy Censorship Principle even if one thinks it is a good moral principle.

The Maximal Autonomy Censor may want to modify their approach if they are worried about paternalism towards individuals. This modified censor may think that paternalism is not justified but that the government is morally entitled to defend the positive autonomy of individuals that is threatened by others. The modified principle is as follows:

The Modified Autonomy Censorship Principle:

Censorship of speech is justified only when the act of censorship serves to protect individuals from losses of positive autonomy on balance due to the speech acts of others.

This principle expresses the view that the government is entitled to use coercion through censorship to maintain the positive autonomy of individuals. This principle extends beyond the moral principle endorsed by the Liberal Censor because the Liberal Censor is concerned about preventing infringements on the negative autonomy of individuals. Adopting such a standard is incredibly problematic because all kinds of speech that we ordinarily think individuals are entitled to express can lead to losses of positive autonomy. Recall in chapter 2 the case of Paul who is deeply passionate about cooking and living life as a chef. If Paul's husband makes a disparaging remark about Paul's cooking then that statement will do enormous damage to Paul's positive autonomy because he will lose faith in himself and be less able to pursue the long term value that he set for himself to be a successful chef. A person who consistently advocates for the Modified Autonomy Censorship Principle must be willing to censor the speech of the husband to criticize Paul's cooking when those critical comments result in losses to Paul's positive autonomy on balance.¹³⁹ But such an approach is absurd because Paul should not have a right to shield his self-esteem through forcible censorship.

There may be obligations that rational agents have to promote the positive autonomy of others through charitable giving, time spent volunteering, Good Samaritan acts, and refraining from speaking on occasions when it would

¹³⁹ In this example assume that the losses of positive autonomy to Paul's husband from being censored are outweighed by the losses of positive autonomy that Paul experiences.

unnecessarily harm the feelings of others or be rude. However, as it was discussed in chapter 2, such obligations should be considered as imperfect duties that do not have to be performed on every occasion. The Maximal Autonomy Censor disagrees and states holds that positive autonomy must be promoted on a consistent basis. This leads to the conclusion that censorship is justified in all occasions when it promotes more positive autonomy on balance. Thus, the Maximal Autonomy Censor turns the promotion of positive autonomy into a perfect duty that must be performed every time. The problem with this approach is that it creates a principle that is too binding and too broad. I think that most readers will believe that the state is not morally justified in censoring Paul's husband from criticizing his cooking and pursuing life as a chef. If one has this assumption in mind then the Modified Autonomy Censorship Principle is substantively lacking because it asserts that censorship is justified in such a case.

Respect for autonomy is the fundamental animating value that is accepted and promoted by the Liberal Censor. Nonetheless, the Liberal Censor sees only violations of the negative autonomy of others relating to the cases considered earlier in this chapter as justifying censorship. The Maximal Autonomy Censor goes too far and asserts an untenable view that paternalism or a more expansive censorship is permissible in the name of promoting positive autonomy so that people are able to realize their central values, plans, and projects in life. While such forms of censorship can be rationally consistent, I have rejected all forms of censorship in the name of positive autonomy as substantively unsound. People that commit to these censorship principle are actually committing themselves to a dangerous vision of state authority. For example, in Plato's *Republic*, an appalling vision of censorship is illustrated where Philosopher Kings or Queens are seen as having the duty to censor and limit freedom in the name of creating a flourishing society. Plato's argument is ultimately one where

the freedom of the individual must be constrained so that positive autonomy on the whole can be promoted where people can live a flourishing life. For instance, Plato says that we must “supervise the storytellers. We’ll select their stories whenever they are fine or beautiful and reject them when they aren’t.”¹⁴⁰ Plato argues for this censorship because he sees it as necessary for cultivating virtue. The virtuous person is ultimately the most free or autonomous person in Plato’s eyes. Plato sees censorship as a means to promote positive autonomy or virtue and that is why he devotes so much time to outlining in unsettling detail how the poets, singers, storytellers, artists, and teachers should be censored to avoid leading the masses away from the path of virtue.¹⁴¹ Moreover, Plato schemes to have the rulers deceive the people with a story about the myth of the metals to manipulate the people in accepting the station that is best for them in life and staying in it.¹⁴²

Plato might well be right. A state that censors the storytellers, propagates lies, and bans the expression of art or music that leads people away from virtue may end up leading to a more virtuous society where people have more positive autonomy on balance. But the gains to autonomy in this manner would be shadow gains because individuals did not have the freedom to completely direct themselves towards the ends they want to pursue in life along the way. The problem with the Maximal Autonomy Censor is that it is based on moral principles that cannot rule out Plato’s idealized society if it truly does promote positive autonomy. The Maximal Autonomy Censor is a consistent censor but it offers a nightmare principle that grants the government a harrowing amount of control over peoples’ lives. A person may think that it is desirable

¹⁴⁰ Plato, *Republic*, trans. G.M.A. Grube (Hackett Publishing Company, 1992), 53.

¹⁴¹ Plato, 71–78.

¹⁴² Plato, 91.

to violate the autonomy of others to promote autonomy overall. This vision does not appeal to me. Individuals must be free to exercise their autonomy so long as they are not violating the rights of others. Individuals will make mistakes along the way and will make it harder for themselves to realize the life they want in life. However, this uneven path is much more preferable than a moral principle that allows censorship in the name of positive autonomy.

Section V: The Truth Censor

The last consistent censor that I will consider is the Truth Censor which is the censor that is focused on censorship as a means of promoting and preserving the truth in the world. When individuals are allowed to speak freely they can spread misinformation, tell lies, and distort reality. Truth Censors are committed to the idea that governments should be empowered to restrict speech that harms the pursuit of truth. Their abiding censorship principle can be expressed as follows:

The Truth Censorship Principle

It is justifiable to censor speech when that speech is false or contributes more on balance to a false understanding of the world.

The Liberal Censor allows censorship for intentional falsehoods in some cases where someone says something intentionally false or with reckless disregard for the truth that harms the reputation of others. However, the Truth Censor goes beyond this standard and also thinks that false statements can be censored even when they do not cause reputational damage to individuals. There are people who claim the Earth is flat, the Holocaust never occurred, or that humans do not contribute to climate change and these people are making false statements. The Truth Censor believes that we can censor these kinds of statements simply *because they are false*

and not because they do things like cause offense and contribute to losses of social utility.

The need to censor truth to more effectively pursue knowledge has been defended within the context of university campuses. Sarah Conly argues that because people suffer from certain cognitive biases and have limited time on college campuses that speech should be censored when it detracts from a university's mission to pursue the truth.¹⁴³ She explains that "Speech should be banned because some speech is not educational: It's false, obviously false, and too obviously false to serve as a useful focus for intellectual criticism."¹⁴⁴ Conly's point is that some speech, like saying the Holocaust never happened, are obviously false statements which waste time and contribute to an inaccurate understanding of the world. The Truth Censor chooses to expand Conly's arguments beyond the scope of college campuses to society as a whole.¹⁴⁵ Speech that is obviously false and contributes to false understandings of the world are liable to be censored because they endanger truth in the world.

An immediate problem with the Truth Censor is that censorship is a process that has perhaps resulted in the most destruction of knowledge and truth in human history. Governments used censorship to ban books that challenged religious orthodoxy, exercised prior restraint to stop the publication of books they disliked, and chilled speech through inquisitions and through executing novel thinkers. Government repression of knowledge is highlighted by Galileo's trial for claiming that the Earth revolves around the Sun. John Stuart Mill wrote his famous defense of

¹⁴³ Sarah Conly, "When Free Speech Is False Speech," in *The Value and Limits of Academic Speech: Philosophical, Political, and Legal Perspectives*, ed. Donald Downs and Chris W. Surprenant (Routledge, 2018), 299–309.

¹⁴⁴ Conly, 299.

¹⁴⁵ I am arguing that the Truth Censor makes this extensional argument but Conly may not actually support the position of the Truth Censor.

freedom of expression as a response to concerns that he had with government censorship and the excessive imposition of governments over individuals. Mill saw censorship as an acute threat to the search for truth. He writes that censorship is dangerous because “If the opinion is right, they are deprived of the opportunity of exchanging error for truth: if wrong, they lose, what is almost as great a benefit, the clearer perception and livelier impression of truth, produced by its collision with error.”¹⁴⁶ Mill argues that censorship risks silencing the truth and that certain problematic arguments can contain valuable portions of truth which should not be suppressed. Furthermore, censoring opinions can lead people to be more dogmatic because they no longer have to examine and present arguments for why they hold certain claims to be true. Mill states that “The fatal tendency of mankind to leave off thinking about a thing when it is no longer doubtful, is the cause of half their errors.”¹⁴⁷

A system of free expression where individuals must present their ideas persuasively to get them accepted by society as a whole has led to a dramatic increase in societal knowledge. Jonathan Rauch documents in his book *Kindly Inquisitors* how the process of liberal science must be safeguarded to protect the cultivation of knowledge and truth in society.¹⁴⁸ Rauch emphasizes that the process of liberal science works because it imposes two binding obligations on society that everyone must be allowed “to err and criticize, even obnoxiously, and to submit everybody’s beliefs – including our own – to public checking before claiming that they deserve to be accepted as knowledge.”¹⁴⁹ The Truth Censor assaults both of these obligations of

¹⁴⁶ Mill, *On Liberty*, 87.

¹⁴⁷ Mill, 110.

¹⁴⁸ Rauch, *Kindly Inquisitors: The New Attacks on Free Thought*.

¹⁴⁹ Rauch, 154.

liberal science because these censors prevent individuals from erring by stating falsehoods and reject the system of public checking of knowledge. Liberal science works because it involves the decentralized clash of ideas but the Truth Censor destroys this process of decentralization by empowering certain centralized government elites to censor false statements. When false ideas are silenced then the scientific process can no longer kick in to refute them or develop more compelling approaches to viewing the world.

The Truth Censor can take two different approaches. The first approach is to side with Mill, Rauch, and the process of liberal science and restrict the government from censoring things just because they are false. This decision acknowledges that free speech and open inquiry are indispensable to the maintenance of truth and the production of knowledge and so it concretely protects freedom of expression. If this road is taken then the Truth Censor practically aligns itself with the Liberal Censor which also says that a statement being false is not a sufficient reason to censor it. The second approach is that the Truth Censor rejects the foundational commitments of liberal science and states that government officials should have the authority to censor certain statements that they believe are false. If this route is taken then there are several practical difficulties that must be surmounted. First, the government must demonstrate that the government's centralized censorship efforts are more effective at promoting truth and knowledge than the decentralized liberal scientific system. Second, the government must outline how its censors are justified to hold their elite status and be capable of consistently discerning fact from fiction to censor speech in a way that is in accordance with the Truth Censorship Principle. Both of these standards need to be demonstrated before we should feel comfortable with the government censoring in the name of truth and these are high argumentative burdens.

Even if these practical objections are solved, we still would not want to accept the Truth Censorship Principle on its face. The reason is that truth has no intrinsic value. Just because you have grasped a true fact about the world does not mean that you have inherently grasped something of value. Suppose that there are currently 50,532,764, apples in the world at the moment. I have no idea if that is remotely the case because the truth of that fact does not personally matter to me. But assume that in this moment you have encountered a true fact of the world. So what? If you are an apple farmer then the fact of how many apples there are in the world may matter to you but that fact would only matter instrumentally in that it would give you a better idea of the world's apple needs. Knowing the truth about certain things is incredibly useful because it can help us act more effectively, realize our ends, or improve our personal welfare. Nonetheless, there are plenty of facts about the world that have no instrumental value. It is not clear why such true facts of the world would suddenly have intrinsic value and be worthy ends in themselves. To highlight this issue consider:

The False Preacher

Jill has a guilty pleasure. She loves saying intentionally false things. To avoid being laughed at, once a week she walks deep into the woods where no one can hear her and shouts towards the heavens all sort of false statements like $2+2=5$ and that the world is flat. She is completely alone and no one can hear her and her false statements have absolutely no impact on anyone else. One day you are aware that Jill is walking into the woods to say false things. You are a committed Truth Censor and think that protecting the truth is in itself a justifying reason for censorship. Thus, you decide to use force to stop Jill from entering the woods and suppress her ability to say false things.

While Jill is certainly an odd character, it appears like she has not done anything that gives you an entitlement to limit her speech by force. Jill's decision to yell falsehoods is an exercise of her freedom and this decision has no negative effects other than the small loss of her own time.¹⁵⁰ The fact that Jill intentionally says $2+2=5$ does not seem to be a justifying reason to censor her and override the exercise of her freedom. Facts do not have inherent value because they are true so there is not an inherent reason to warrant censorship.

A main argument of this chapter is that all of the consistent censors are fundamentally opposed to each other in their foundational commitments. A consistent Truth Censor is opposed to view of the Liberal Censor because it is committed to censoring people who say false things even when they do not violate the rights of others. This is seen in the false preacher case where Jill exercises her freedom to say false things in the woods without violating the rights of others. The Liberal Censor would not censor Jill but the Truth Censor would. The Truth Censor is also opposed to utilitarianism because the truth can be very painful sometimes and lead to losses in utility. Suppressing the truth in certain cases would help maximize utility but violate the basic commitments of the Truth Censor. Stating falsehoods could also help the Radical Egalitarian Censor achieve their aims of promoting equality of outcomes between persons. Finally, protecting the truth runs contrary of the aims of the Maximal Autonomy Censor to maximize positive autonomy. Recall Plato's myth of the metals where the city's leaders tell lies to the population so that the citizens are more inclined to accept the best position they should be in life.¹⁵¹ Telling the Myth of the

¹⁵⁰ Assume that in this case Jill's choice to yell these falsehoods in the woods does not make her more likely to say false things in other contexts.

¹⁵¹ Plato, *Republic*, 91.

Metals would violate the principles the Truth Censor is committed to but could end up enhancing the positive autonomy of people overall. Thus, one cannot be a principled Truth Censor while also committing to any of the other consistent censors.

Conclusion

The focus of this project has been to critically examine and reflect on the moral foundations of various proposed systems of free expression. This systematic analysis also includes a framework for evaluating morally justifiable forms of censorship. Applying this framework, my central contention has been that the liberal model of censorship is the most plausible censorship model. This model says that censorship of speech is only morally justified when speech violates the autonomy rights of another. Other models of censorship were evaluated and rejected for either being methodologically inconsistent or for being morally implausible. This thesis has been divided into three sections. The first section looks at the normative grounds of moral obligation and the methodological requirements necessary for constructing a valid moral principle. The second section then applies these methodological requirements to four proposed models of censorship and found that they failed to advance valid moral principles. The final section then looks at five models of censorship with valid moral principles and draws attention to their substantive implications.

The first part of this thesis looked at the source of normativity for why humans have moral obligations. I argued that moral obligations flow from the need for rational agents to deliberate and make choices in a manner that is intelligible so that their agency can be maintained over time. My argument was that

1. Humans are agents
2. Agents ought to live intelligible lives

3. Rationality is the mechanism by which one's life and reality becomes intelligible
4. Therefore, humans ought to be rational

If moral obligation is grounded in rationality then the structure of our moral obligations must align with the structure of rationality. I argued that the structure of rationality must be based in universally coherent principles. The first standard of universality is that the objective state of intelligible reality must be true for all rational agents. The second standard of coherence is that all valid principles must cohere and cannot contradict each other. These standards impose a methodological framework on any valid moral principle that rational agents have a normative obligation to comply with.

The duty to be rational implies a duty to avoid hypocrisy. A hypocrite is someone whose actions do not cohere because they do not act in a manner that can be universalized by applying unequal standards across different rational agents. Whenever someone subjects another to sanction on the basis of a universally coherent moral principle without subjecting themselves to the same standard then they are behaving hypocritically and irrationally. I then applied this framework to evaluate the general moral wrong of lying and censorship. Lying and censorship are generally wrong because they fail to respect the equal moral standing of others due to the application of unequal standards that prioritizes the interests of some over others on the basis of rationally incoherent standards. I point out that a system of free and truthful communication is of enormous value to the vast majority of rational agents because it

- Allows us to collectively work with others to realize our moral aims
- Provides us access into the minds of others to assign moral responsibility

- Enables individuals to cultivate their autonomy and individual identity
- Preserves the ability of agents to fulfill their duties to act authentically

Thus, if rational agents practically commit themselves to the value of communication they would be acting irrationally if they engage in actions that subverts this very same system of communication.

There are cases where there are principled justifications for lying and for censoring the speech of others. I argue that the best principled justification for these actions rest in committing to the value of autonomy. Any viable system of communication should preserve the equal freedom of rational agents in order to be universally coherent. Actors who wrongly infringe on the autonomy of another rational agent make them liable to being deceived or censored if that is what is necessary to protect or redress the violation of autonomy and return the state of affairs to one of equal freedom. This argument forms the basis of the liberal censorship model which holds that censorship is justified only against individuals who violate the rights rational agents have to independence as a sovereign agent.

The second section applies the methodological framework of valid moral principles to four censorship models that present themselves as types of liberal censors. These censors want to allow speech to be broadly expressed and do not want to empower the state with an extensive amount of censorship powers. However, these censors each claim that the state should be allowed to censor speech that violates cultural standards, offensive speech, hate speech, and speech that undermines democratic discourse. I went into detail explaining that if these principles were to be applied to censor speech in a non-arbitrary way then they would be illiberal principles that would allow for the censorship of a lot of behavior and speech that would be

protected under a liberal regime. I labeled these censors as hypocritical because they hold certain speech to different standards than other speech that are covered under the same principle. To resolve the hypocrisy, these censors must either reject the illiberal censorship principles they advance or no longer commit to a liberal system of free expression.

The final section of this thesis looks into the implications of departing from a liberal system of free expression. I begin by developing in detail what this system looks like. The guiding value of the Liberal Censor is the value of autonomy and this censor thinks the only justifying reason for censorship is maintaining the independence of rational agents according to a state of equal freedom. The following forms of speech violate the rights of individuals by infringing on the sphere of independence people are entitled to as rational agents:

- Incitement which provokes the immediate and foreseeable violations of an individual's bodily or property rights such as Theater Fire Case.
- Defamation which damages a person's entitlement to having their reputation not be destroyed without a fair hearing due to false or reckless speech.
- Fraud which violates the property rights of individuals to consent to property transfers without being subjected to misrepresentations.
- Harassment which violates the rights of individuals to be free from certain kinds of severe and pervasive forms of interference in a sphere of control they are entitled to exercise control over.
- Perjury which violates a person's entitlement to not be subjected to coercive force without a fair hearing and due process.

- Risky Disclosure which violates the rights of individuals not to be exposed to an unreasonable risk of harm such as in the Nuclear Codes Case.
- Assaultive Speech which violates rights by causing non-belief-mediated distress such as screaming on a residential street late at night

I then consider other consistent censors who reject the liberal model of free expression. These censors utilize universally coherent moral principles so they are immune from methodological objections. Nonetheless, committing to these consistent censors has extreme substantive and practical implications. I argue that the Utilitarian, Radical Egalitarian, Maximal Autonomy, and Truth Censors are all not plausible censors because they commit themselves to governing moral values which are not inherently good values to commit to. Moreover, it is very dangerous to allow governments comprised of fallible state officials to censor along the lines of these principles because of the danger of improper censorship which ends up undermining the source of the principle itself. I also argued that these consistent censors cannot be simultaneously committed to because their foundational values are in irreconcilable tension.

These arguments have several important implications for leadership. Leaders are ultimately rational agents so they will have the same binding moral obligations grounded in rationality as everyone else. Leaders do not get to be an exception from the requirements that they act rationally in ways that are universally coherent. Thus, leaders are not morally entitled to act hypocritically. Rational considerations impose constraints on the moral permissibility of the coercive actions leaders can engage in. Institutional officials will be in positions to shape and enforce the law. They have the responsibility to do so in a consistent manner. They should also act consistently in a

way that is plausible and when their aims will not be practically counterproductive. This framework implies that leaders should only censor speech according to the liberal censor model. If the foregoing arguments are successful then leaders should not justify censorship of speech for cultural standards, offense, hate speech, effective civic discourse, utility, equality, human flourishing, or truth.

Being moral may at times be unsettling or inconvenient. It may challenge our deepest impulses and demand more of us than we want to give. However, the demand to be moral is a responsibility that never leaves us. It is something that is always the right thing to do. As rational agents we live a life of dignity when we act rationally and when we pursue this moral path. As Ronald Dworkin once said, “Without dignity our lives are only blinks of duration. But if we manage to lead a good life well, we create something more. We write a subscript to our mortality. We make our lives tiny diamonds in the cosmic sands.”¹⁵² We should never stop speaking as we write out this dignified subscript to our lives and we must defend the freedom of others to follow this same path.

¹⁵² Ronald Dworkin, *Justice for Hedgehogs* (Harvard University Press, 2011), 423.

Bibliography

- Altman, Andrew. "Liberalism and Campus Hate Speech: A Philosophical Examination." *Ethics* 103, no. 2 (1993): 302–17.
- Ash, Timothy Garton. *Free Speech: Ten Principles for a Connected World*. Yale University Press, 2016.
- Berlin, Isaiah. *Four Essays on Liberty*. Oxford University Press, 1969.
- Cohen, G. A. "Robert Nozick and Wilt Chamberlain: How Patterns Preserve Liberty." *Erkenntnis* (1975-) 11, no. 1 (1977): 5–23.
- Cohen, G.A. *On the Currency of Egalitarian Justice and Other Essays in Political Philosophy*. Princeton University Press, 2011.
- Justia Law. "Cohen v. California, 403 U.S. 15 (1971)." Accessed February 28, 2021. <https://supreme.justia.com/cases/federal/us/403/15/>.
- Conly, Sarah. *Against Autonomy: Justifying Coercive Paternalism*. Cambridge University Press, 2013.
- . "When Free Speech Is False Speech." In *The Value and Limits of Academic Speech: Philosophical, Political, and Legal Perspectives*, edited by Donald Downs and Chris W. Surprenant, 299–309. Routledge, 2018.
- Cosse, Eva. "The Alarming Rise of Anti-Semitism in Europe." Human Rights Watch, June 4, 2019. <https://www.hrw.org/news/2019/06/04/alarming-rise-anti-semitism-europe>.
- Dworkin, Ronald. *Justice for Hedgehogs*. Harvard University Press, 2011.
- Feinberg, Joel. *The Moral Limits of the Criminal Law: Offense to Others*. Oxford University Press, 1988.
- Feldman, Richard. "The Ethics of Belief." *Philosophy and Phenomenological Research* 60, no. 3 (2000): 667–95.

- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* 68, no. 1 (1971): 5–20.
- Goodreads. "Freedom Of Speech Quotes." Accessed February 28, 2021.
<https://www.goodreads.com/quotes/tag/freedom-of-speech>.
- Freiman, Christopher. *Unequivocal Justice*. Routledge, 2019.
- Gasset, Jose Ortega. *The Revolt of the Masses*. W.W. Norton & Company, 1994.
- Hamilton, Alexander, James Madison, and John Jay. *The Federalist Papers*. Edited by Clinton Rossiter. New American Library, 2003.
- Hare, Ivan, and James Weinstein. *Extreme Speech and Democracy*. 1 edition. Oxford ; New York: Oxford University Press, 2009.
- Herman, Barbara. *The Practice of Moral Judgment*. Harvard University Press, 1993.
- Hill, Thomas. "Pains and Projects: Justifying to Oneself." In *Autonomy and Self-Respect*, 173–88. Cambridge University Press, 1991.
- Howard, Jeffrey W. "Dangerous Speech." *Philosophy & Public Affairs* 47, no. 2 (2019): 208–54. <https://doi.org/10.1111/papa.12145>.
- Huemer, Michael. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. Palgrave Macmillan, 2013.
- Kant, Immanuel. *Groundwork of the Metaphysic of Morals*. Translated by Thomas Abbott. Digireads.com Publishing, 2017.
- Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge University Press, 1996.
- . "The Right to Lie: Kant on Dealing with Evil." *Philosophy and Public Affairs* 15, no. 4 (1986): 325–49.
- . *The Sources of Normativity*. Cambridge University Press, 1996.

- Langton, Rae. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* 22, no. 4 (1993): 293–330.
- Lillehammer, Hallyard. "The Importance of Being Rational." *Notre Dame Philosophical Reviews*, April 2019. <https://ndpr.nd.edu/news/the-importance-of-being-rational/>.
- Lippert-Rasmussen, Kasper. *Relational Egalitarianism: Living as Equals*. Cambridge, United Kingdom: Cambridge University Press, 2018.
- Machan, Tibor. *Liberty and Equality*. Hoover Press, 2013.
- Mill, John Stuart. *On Liberty*. Edited by David Bromwich and George Kateb. Yale University Press, 2003.
- . *Utilitarianism and On Liberty*. Edited by Mary Warnock. 2nd ed. John Wiley & Sons, 2003.
- Nielsen, Kai. *Equality and Liberty: A Defense of Radical Egalitarianism*. Rowman & Littlefield Publishers, 1985.
- Nozick, Robert. *Anarchy, State, and Utopia*. Basic Books, 2013.
- Parfit, Derek. *Reasons and Persons*. Oxford University Press, 1986.
- Plato. *Republic*. Translated by G.M.A. Grube. Hackett Publishing Company, 1992.
- Price, Terry. *Leadership Ethics: An Introduction*. Cambridge University Press, 2008.
- Quinn, Ben. "Hate Crimes Double in Five Years in England and Wales." *The Guardian*, October 15, 2019. <https://www.theguardian.com/society/2019/oct/15/hate-crimes-double-england-wales>.
- Rachels, James. "The Challenge of Cultural Relativism." In *Exploring Philosophy: An Introductory Anthology*, edited by S.M. Cahn, 4th ed., 1–12. Oxford University Press, 2009.

- Rauch, Jonathan. *Kindly Inquisitors: The New Attacks on Free Thought*. 2nd ed. The University of Chicago Press, 2013.
- Rawls, John. *Justice As Fairness*. Harvard University Press, 2001.
- Rinard, Susanna. "Equal Treatment for Belief." *Philosophical Studies* 176, no. 7 (July 1, 2019): 1923–50. <https://doi.org/10.1007/s11098-018-1104-9>.
- Ripstein, Arthur. *Force and Freedom: Kant's Legal and Political Philosophy*. Harvard University Press, 2010.
- Savage, Rachel. "Rising Populism Stokes Homophobic Hate Speech across Europe." Reuters, February 4, 2020. <https://www.reuters.com/article/us-europe-lgbt-rights-trfn/rising-populism-stokes-homophobic-hate-speech-across-europe-rights-group-idUSKBN1ZY0X3>.
- Scanlon, T.M. *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge University Press, 2003.
- Schapiro, Tamar. "Kantian Rigorism and Mitigating Circumstances." *Ethics* 117, no. 1 (2006): 32–57.
- Shiffrin, Seana. *Speech Matters: On Lying, Morality, and the Law*. Princeton University Press, 2014.
- Singer, Peter. "Famine, Affluence, and Morality." *Philosophy & Public Affairs* 1, no. 3 (1972): 229–43.
- Smart, J.C.C., and Bernard Williams. *Utilitarianism For & Against*. Cambridge University Press, 1973.
- Strossen, Nadine. *Hate: Why We Should Resist It with Free Speech, Not Censorship*. Oxford University Press, 2020.
- Sunstein, Cass. *Democracy and the Problem of Free Speech*. Free Press, 1995.
- Thomson, Judith. *The Realm of Rights*. Harvard University Press, 1990.

- Vonnegut, Kurt. "Harrison Bergeron." In *Philosophy, Politics, and Economics: An Anthology*, edited by Jonathan Anomaly, Geoffrey Brennan, Michael Munger, and Geoffrey Sayre-McCord. Oxford University Press, 2016.
- Waldron, Jeremy. *The Harm in Hate Speech*. Harvard University Press, 2014.
- Wallace, Jay. "Hypocrisy, Moral Address, and the Equal Standing of Persons." *Philosophy and Public Affairs* 38, no. 4 (2010): 307–41.
- Watson, Gary. "Free Agency." *The Journal of Philosophy* 72, no. 8 (1975): 205–20.
- Justia Law. "West Virginia State Bd. of Educ. v. Barnette, 319 U.S. 624 (1943)." Accessed April 11, 2021.
<https://supreme.justia.com/cases/federal/us/319/624/>.
- Justia Law. "Whitney v. California, 274 U.S. 357 (1927)." Accessed December 27, 2020. <https://supreme.justia.com/cases/federal/us/274/357/>.